

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 June 2001 (07.06.2001)

PCT

(10) International Publication Number
WO 01/40466 A2

(51) International Patent Classification ⁷ :	C12N 15/12, C07K 14/47, 14/705, C12N 15/62, C07K 16/18, 16/28, G01N 33/53, A61K 38/17, C12Q 1/68	PCT/US99/30999	20 December 1999 (20.12.1999)	US
		PCT/US99/31243	30 December 1999 (30.12.1999)	US
(21) International Application Number:	PCT/US00/32678	PCT/US99/31274	30 December 1999 (30.12.1999)	US
(22) International Filing Date:	1 December 2000 (01.12.2000)	PCT/US00/00219	5 January 2000 (05.01.2000)	US
		PCT/US00/00277	6 January 2000 (06.01.2000)	US
(25) Filing Language:	English	PCT/US00/00376	6 January 2000 (06.01.2000)	US
		PCT/US00/03565		
(26) Publication Language:	English	PCT/US00/04341	11 February 2000 (11.02.2000)	US
(30) Priority Data:		PCT/US00/04342	18 February 2000 (18.02.2000)	US
PCT/US99/28301	1 December 1999 (01.12.1999)	US	18 February 2000 (18.02.2000)	US
PCT/US99/28634	1 December 1999 (01.12.1999)	US	22 February 2000 (22.02.2000)	US
PCT/US99/28551	2 December 1999 (02.12.1999)	US	24 February 2000 (24.02.2000)	US
PCT/US99/28564	2 December 1999 (02.12.1999)	US	24 February 2000 (24.02.2000)	US
PCT/US99/28565	2 December 1999 (02.12.1999)	US	1 March 2000 (01.03.2000)	US
60/170,262	9 December 1999 (09.12.1999)	US	2 March 2000 (02.03.2000)	US
PCT/US99/30095	16 December 1999 (16.12.1999)	US	3 March 2000 (03.03.2000)	US
PCT/US99/30911	20 December 1999 (20.12.1999)	US	10 March 2000 (10.03.2000)	US
			15 March 2000 (15.03.2000)	US
			20 March 2000 (20.03.2000)	US
			21 March 2000 (21.03.2000)	US

[Continued on next page]

(54) Title: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

MSTMFADTLLIVFISVCTALLAEGITWVLVYRTDKYKRLKAEVEKQSKKLEKKKETITESAGR
QQKKKIERQEEKLKNNRDLMSVRMKSMAIGFCFTALMGMFNSIFDGRVVAKLPFTPLSYIQ
GLSHRNLLGDDTDCSFIFLYILCTMSIRQNIQKILGLAPSRAATKQAGGFLGPPPPSGKFS

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 103-109, 163-169

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 53-57

(57) Abstract: The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

Best Available Copy

WO 01/40466 A2



- | | | | |
|----------------|--------------------------------|----|--|
| PCT/US00/08439 | 30 March 2000 (30.03.2000) | US | Los Altos, CA 94024 (US). SMITH, Victoria [AU/US]; |
| PCT/US00/13705 | 17 May 2000 (17.05.2000) | US | 19 Dwight Road, Burlingame, CA 94010 (US). STEW- |
| PCT/US00/14042 | 22 May 2000 (22.05.2000) | US | ART, Timothy, A. [US/US]; 465 Douglass Street, San |
| PCT/US00/14941 | 30 May 2000 (30.05.2000) | US | Francisco, CA 94114 (US). TUMAS, Daniel [US/US]; |
| PCT/US00/15264 | 2 June 2000 (02.06.2000) | US | 3 Rae Avenue, Orinda, CA 94563 (US). WATANABE, |
| 60/209,832 | 5 June 2000 (05.06.2000) | US | Colin, K. [US/US]; 128 Corliss Drive, Moraga, CA 94556 |
| PCT/US00/20710 | 28 July 2000 (28.07.2000) | US | (US). WOOD, William, I. [US/US]; 35 Southdown Court, |
| PCT/US00/22031 | 11 August 2000 (11.08.2000) | US | Hillsborough, CA 94010 (US). ZHANG, Zemin [CN/US]; |
| PCT/US00/23522 | 23 August 2000 (23.08.2000) | US | 876 Taurus Drive, Foster City, CA 94404 (US). |
| PCT/US00/23328 | 24 August 2000 (24.08.2000) | US | |
| 60/000,000 | 15 September 2000 (15.09.2000) | US | |
| PCT/US00/30952 | 8 November 2000 (08.11.2000) | US | (74) Agents: KRESNAK, Mark, T. et al.; Genentech, Inc., |
| PCT/US00/30873 | 10 November 2000 (10.11.2000) | US | MS49, 1 DNA Way, South San Francisco, CA 94080-4990 |
| | | | (US). |
- (71) Applicant (for all designated States except US): **GENEN-TECH, INC.** [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **BAKER, Kevin, P.** [GB/US]; 14006 Indian Run Drive, Darnestown, MD 20878 (US). **BERESINI, Maureen** [US/US]; 611 Stetson Street, Moss Beach, CA 94038 (US). **DEFORGE, Laura** [US/US]; 1175 Manzanita Drive, Pacifica, CA 94044 (US). **DESNOYERS, Luc** [CA/US]; 2050 Stockton Street, San Francisco, CA 94133 (US). **FILVAROFF, Ellen** [US/US]; 538 18th Avenue, San Francisco, CA 94121 (US). **GAO, Wei-Qiang** [CN/US]; 641 Pilgrim Drive, Foster City, CA 94404 (US). **GERRITSEN, Mary, E.** [CA/US]; 541 Parrott Drive, San Mateo, CA 94402 (US). **GODDARD, Audrey** [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). **GODOWSKI, Paul, J.** [US/US]; 2627 Easton Drive, Burlingame, CA 94010 (US). **GURNEY, Austin, L.** [US/US]; 1 Debbie Lane, Belmont, CA 94002 (US). **SHERWOOD, Steven** [US/US]; 995 Lundy Lane,
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:**
— Without international search report and to be republished upon receipt of that report.
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

FIELD OF THE INVENTION

5 The present invention relates generally to the identification and isolation of novel DNA and to the recombinant production of novel polypeptides.

BACKGROUND OF THE INVENTION

10 Extracellular proteins play important roles in, among other things, the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. These secreted polypeptides or signaling molecules normally pass through the cellular secretory pathway to reach their site of
15 action in the extracellular environment.

Secreted proteins have various industrial applications, including as pharmaceuticals, diagnostics, biosensors and bioreactors. Most protein drugs available at present, such as thrombolytic agents, interferons, interleukins, erythropoietins, colony stimulating factors, and various other cytokines, are secretory proteins. Their receptors, which are membrane proteins, also have potential as therapeutic or diagnostic agents. Efforts
20 are being undertaken by both industry and academia to identify new, native secreted proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci. 93:7108-7113 (1996); U.S. Patent No. 5,536,637].

Membrane-bound proteins and receptors can play important roles in, among other things, the formation,
25 differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins.
30 Such membrane-bound proteins and cell receptors include, but are not limited to, cytokine receptors, receptor kinases, receptor phosphatases, receptors involved in cell-cell interactions, and cellular adhesion molecules like selectins and integrins. For instance, transduction of signals that regulate cell growth and differentiation is regulated in part by phosphorylation of various cellular proteins. Protein tyrosine kinases, enzymes that catalyze that process, can also act as growth factor receptors. Examples include fibroblast growth factor receptor and

nerve growth factor receptor.

Membrane-bound proteins and receptor molecules have various industrial applications, including as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be employed as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction.

- 5 Efforts are being undertaken by both industry and academia to identify new, native receptor or membrane-bound proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor or membrane-bound proteins.

SUMMARY OF THE INVENTION

- 10 In one embodiment, the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence that encodes a PRO polypeptide.

- 15 In one aspect, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule encoding a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

- 20 In other aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94%
- 25 30 35

nucleic acid sequence identity, alternatively at least about 95 % nucleic acid sequence identity, alternatively at least about 96 % nucleic acid sequence identity, alternatively at least about 97 % nucleic acid sequence identity, alternatively at least about 98 % nucleic acid sequence identity and alternatively at least about 99 % nucleic acid sequence identity to (a) a DNA molecule comprising the coding sequence of a full-length PRO polypeptide cDNA as disclosed herein, the coding sequence of a PRO polypeptide lacking the signal peptide as disclosed herein, the coding sequence of an extracellular domain of a transmembrane PRO polypeptide, with or without the signal peptide, as disclosed herein or the coding sequence of any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In a further aspect, the invention concerns an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80 % nucleic acid sequence identity, alternatively at least about 81 % nucleic acid sequence identity, alternatively at least about 82 % nucleic acid sequence identity, alternatively at least about 83 % nucleic acid sequence identity, alternatively at least about 84 % nucleic acid sequence identity, alternatively at least about 85 % nucleic acid sequence identity, alternatively at least about 86 % nucleic acid sequence identity, alternatively at least about 87 % nucleic acid sequence identity, alternatively at least about 88 % nucleic acid sequence identity, alternatively at least about 89 % nucleic acid sequence identity, alternatively at least about 90 % nucleic acid sequence identity, alternatively at least about 91 % nucleic acid sequence identity, alternatively at least about 92 % nucleic acid sequence identity, alternatively at least about 93 % nucleic acid sequence identity, alternatively at least about 94 % nucleic acid sequence identity, alternatively at least about 95 % nucleic acid sequence identity, alternatively at least about 96 % nucleic acid sequence identity, alternatively at least about 97 % nucleic acid sequence identity, alternatively at least about 98 % nucleic acid sequence identity and alternatively at least about 99 % nucleic acid sequence identity to (a) a DNA molecule that encodes the same mature polypeptide encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein, or (b) the complement of the DNA molecule of (a).

Another aspect the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain(s) of such polypeptide are disclosed herein. Therefore, soluble extracellular domains of the herein described PRO polypeptides are contemplated.

Another embodiment is directed to fragments of a PRO polypeptide coding sequence, or the complement thereof, that may find use as, for example, hybridization probes, for encoding fragments of a PRO polypeptide that may optionally encode a polypeptide comprising a binding site for an anti-PRO antibody or as antisense oligonucleotide probes. Such nucleic acid fragments are usually at least about 10 nucleotides in length, alternatively at least about 15 nucleotides in length, alternatively at least about 20 nucleotides in length, alternatively at least about 30 nucleotides in length, alternatively at least about 40 nucleotides in length, alternatively at least about 50 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 70 nucleotides in length, alternatively at least about 80 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 100 nucleotides in length, alternatively at least about 110 nucleotides in length, alternatively at least about 120 nucleotides in length,

alternatively at least about 130 nucleotides in length, alternatively at least about 140 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 160 nucleotides in length, alternatively at least about 170 nucleotides in length, alternatively at least about 180 nucleotides in length, alternatively at least about 190 nucleotides in length, alternatively at least about 200 nucleotides in length, alternatively at least about 250 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 350 nucleotides in length, alternatively at least about 400 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 500 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 700 nucleotides in length, alternatively at least about 800 nucleotides in length, alternatively at least about 900 nucleotides in length and alternatively at least about 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length. It is noted that novel fragments of a PRO polypeptide-encoding nucleotide sequence may be determined in a routine manner by aligning the PRO polypeptide-encoding nucleotide sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which PRO polypeptide-encoding nucleotide sequence fragment(s) are novel. All of such PRO polypeptide-encoding nucleotide sequences are contemplated herein. Also contemplated are the PRO polypeptide fragments encoded by these nucleotide molecule fragments, preferably those PRO polypeptide fragments that comprise a binding site for an anti-PRO antibody.

In another embodiment, the invention provides isolated PRO polypeptide encoded by any of the isolated nucleic acid sequences hereinabove identified.

In a certain aspect, the invention concerns an isolated PRO polypeptide, comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein.

In a further aspect, the invention concerns an isolated PRO polypeptide comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid

sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to an amino acid sequence encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein.

In a specific aspect, the invention provides an isolated PRO polypeptide without the N-terminal signal sequence and/or the initiating methionine and is encoded by a nucleotide sequence that encodes such an amino acid sequence as hereinbefore described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

Another aspect the invention provides an isolated PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

In yet another embodiment, the invention concerns agonists and antagonists of a native PRO polypeptide as defined herein. In a particular embodiment, the agonist or antagonist is an anti-PRO antibody or a small molecule.

In a further embodiment, the invention concerns a method of identifying agonists or antagonists to a PRO polypeptide which comprise contacting the PRO polypeptide with a candidate molecule and monitoring a biological activity mediated by said PRO polypeptide. Preferably, the PRO polypeptide is a native PRO polypeptide.

In a still further embodiment, the invention concerns a composition of matter comprising a PRO polypeptide, or an agonist or antagonist of a PRO polypeptide as herein described, or an anti-PRO antibody, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

Another embodiment of the present invention is directed to the use of a PRO polypeptide, or an agonist or antagonist thereof as hereinbefore described, or an anti-PRO antibody, for the preparation of a medicament useful in the treatment of a condition which is responsive to the PRO polypeptide, an agonist or antagonist thereof or an anti-PRO antibody.

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described polypeptides. Host cell comprising any such vector are also provided. By way of example, the host cells may be CHO cells, *E. coli*, or yeast. A process for producing any of the herein described polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired polypeptide and recovering the desired polypeptide from the cell culture.

5 In other embodiments, the invention provides chimeric molecules comprising any of the herein described polypeptides fused to a heterologous polypeptide or amino acid sequence. Example of such chimeric molecules comprise any of the herein described polypeptides fused to an epitope tag sequence or a Fc region of an immunoglobulin.

10 In another embodiment, the invention provides an antibody which binds, preferably specifically, to any of the above or below described polypeptides. Optionally, the antibody is a monoclonal antibody, humanized antibody, antibody fragment or single-chain antibody.

In yet other embodiments, the invention provides oligonucleotide probes which may be useful for isolating genomic and cDNA nucleotide sequences, measuring or detecting expression of an associated gene or as antisense probes, wherein those probes may be derived from any of the above or below described nucleotide sequences. Preferred probe lengths are described above.

15 In yet other embodiments, the present invention is directed to methods of using the PRO polypeptides of the present invention for a variety of uses based upon the functional biological assay data presented in the Examples below.

20 BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a nucleotide sequence (SEQ ID NO:1) of a native sequence PRO177 cDNA, wherein SEQ ID NO:1 is a clone designated herein as "DNA16438-1387".

Figure 2 shows the amino acid sequence (SEQ ID NO:2) derived from the coding sequence of SEQ ID NO:1 shown in Figure 1.

25 Figure 3 shows a nucleotide sequence (SEQ ID NO:3) of a native sequence PRO3574 cDNA, wherein SEQ ID NO:3 is a clone designated herein as "DNA19360-2552".

Figure 4 shows the amino acid sequence (SEQ ID NO:4) derived from the coding sequence of SEQ ID NO:3 shown in Figure 3.

30 Figure 5 shows a nucleotide sequence (SEQ ID NO:5) of a native sequence PRO1280 cDNA, wherein SEQ ID NO:5 is a clone designated herein as "DNA33455-1548".

Figure 6 shows the amino acid sequence (SEQ ID NO:6) derived from the coding sequence of SEQ ID NO:5 shown in Figure 5.

Figure 7 shows a nucleotide sequence (SEQ ID NO:7) of a native sequence PRO4984 cDNA, wherein SEQ ID NO:7 is a clone designated herein as "DNA37155-2651".

35 Figure 8 shows the amino acid sequence (SEQ ID NO:8) derived from the coding sequence of SEQ ID NO:7 shown in Figure 7.

Figure 9 shows a nucleotide sequence (SEQ ID NO:9) of a native sequence PRO4988 cDNA, wherein SEQ ID NO:9 is a clone designated herein as "DNA38269-2654".

Figure 10 shows the amino acid sequence (SEQ ID NO:10) derived from the coding sequence of SEQ ID NO:9 shown in Figure 9.

5 Figure 11 shows a nucleotide sequence (SEQ ID NO:11) of a native sequence PRO305 cDNA, wherein SEQ ID NO:11 is a clone designated herein as "DNA40619-1220".

Figure 12 shows the amino acid sequence (SEQ ID NO:12) derived from the coding sequence of SEQ ID NO:11 shown in Figure 11.

Figure 13 shows a nucleotide sequence (SEQ ID NO:13) of a native sequence PRO1866 cDNA, wherein SEQ ID NO:13 is a clone designated herein as "DNA44174-2513".

10 Figure 14 shows the amino acid sequence (SEQ ID NO:14) derived from the coding sequence of SEQ ID NO:13 shown in Figure 13.

Figure 15 shows a nucleotide sequence (SEQ ID NO:15) of a native sequence PRO4996 cDNA, wherein SEQ ID NO:15 is a clone designated herein as "DNA44675-2662".

15 Figure 16 shows the amino acid sequence (SEQ ID NO:16) derived from the coding sequence of SEQ ID NO:15 shown in Figure 15.

Figure 17 shows a nucleotide sequence (SEQ ID NO:17) of a native sequence PRO4406 cDNA, wherein SEQ ID NO:17 is a clone designated herein as "DNA45408-2615".

Figure 18 shows the amino acid sequence (SEQ ID NO:18) derived from the coding sequence of SEQ ID NO:17 shown in Figure 17.

20 Figure 19 shows a nucleotide sequence (SEQ ID NO:19) of a native sequence PRO1120 cDNA, wherein SEQ ID NO:19 is a clone designated herein as "DNA48606-1479".

Figure 20 shows the amino acid sequence (SEQ ID NO:20) derived from the coding sequence of SEQ ID NO:19 shown in Figure 19.

25 Figure 21 shows a nucleotide sequence (SEQ ID NO:21) of a native sequence PRO4990 cDNA, wherein SEQ ID NO:21 is a clone designated herein as "DNA52753-2656".

Figure 22 shows the amino acid sequence (SEQ ID NO:22) derived from the coding sequence of SEQ ID NO:21 shown in Figure 21.

Figure 23 shows a nucleotide sequence (SEQ ID NO:23) of a native sequence PRO738 cDNA, wherein SEQ ID NO:23 is a clone designated herein as "DNA53915-1258".

30 Figure 24 shows the amino acid sequence (SEQ ID NO:24) derived from the coding sequence of SEQ ID NO:23 shown in Figure 23.

Figure 25 shows a nucleotide sequence (SEQ ID NO:25) of a native sequence PRO3577 cDNA, wherein SEQ ID NO:25 is a clone designated herein as "DNA53991-2553".

35 Figure 26 shows the amino acid sequence (SEQ ID NO:26) derived from the coding sequence of SEQ ID NO:25 shown in Figure 25.

Figure 27 shows a nucleotide sequence (SEQ ID NO:27) of a native sequence PRO1879 cDNA, wherein SEQ ID NO:27 is a clone designated herein as "DNA54009-2517".

Figure 28 shows the amino acid sequence (SEQ ID NO:28) derived from the coding sequence of SEQ ID NO:27 shown in Figure 27.

Figure 29 shows a nucleotide sequence (SEQ ID NO:29) of a native sequence PRO1471 cDNA, wherein SEQ ID NO:29 is a clone designated herein as "DNA56055-1643".

5 Figure 30 shows the amino acid sequence (SEQ ID NO:30) derived from the coding sequence of SEQ ID NO:29 shown in Figure 29.

Figure 31 shows a nucleotide sequence (SEQ ID NO:31) of a native sequence PRO1114 cDNA, wherein SEQ ID NO:31 is a clone designated herein as "DNA57033-1403".

Figure 32 shows the amino acid sequence (SEQ ID NO:32) derived from the coding sequence of SEQ ID NO:31 shown in Figure 31.

10 Figure 33 shows a nucleotide sequence (SEQ ID NO:33) of a native sequence PRO1076 cDNA, wherein SEQ ID NO:33 is a clone designated herein as "DNA57252-1453".

Figure 34 shows the amino acid sequence (SEQ ID NO:34) derived from the coding sequence of SEQ ID NO:33 shown in Figure 33.

15 Figure 35 shows a nucleotide sequence (SEQ ID NO:35) of a native sequence PRO1483 cDNA, wherein SEQ ID NO:35 is a clone designated herein as "DNA58799-1652".

Figure 36 shows the amino acid sequence (SEQ ID NO:36) derived from the coding sequence of SEQ ID NO:35 shown in Figure 35.

Figure 37 shows a nucleotide sequence (SEQ ID NO:37) of a native sequence PRO4985 cDNA, wherein SEQ ID NO:37 is a clone designated herein as "DNA59770-2652".

20 Figure 38 shows the amino acid sequence (SEQ ID NO:38) derived from the coding sequence of SEQ ID NO:37 shown in Figure 37.

Figure 39 shows a nucleotide sequence (SEQ ID NO:39) of a native sequence PRO5000 cDNA, wherein SEQ ID NO:39 is a clone designated herein as "DNA59774-2665".

25 Figure 40 shows the amino acid sequence (SEQ ID NO:40) derived from the coding sequence of SEQ ID NO:39 shown in Figure 39.

Figure 41 shows a nucleotide sequence (SEQ ID NO:41) of a native sequence PRO1881 cDNA, wherein SEQ ID NO:41 is a clone designated herein as "DNA60281-2518".

Figure 42 shows the amino acid sequence (SEQ ID NO:42) derived from the coding sequence of SEQ ID NO:41 shown in Figure 41.

30 Figure 43 shows a nucleotide sequence (SEQ ID NO:43) of a native sequence PRO4314 cDNA, wherein SEQ ID NO:43 is a clone designated herein as "DNA60736-2559".

Figure 44 shows the amino acid sequence (SEQ ID NO:44) derived from the coding sequence of SEQ ID NO:43 shown in Figure 43.

35 Figure 45 shows a nucleotide sequence (SEQ ID NO:45) of a native sequence PRO4987 cDNA, wherein SEQ ID NO:45 is a clone designated herein as "DNA61875-2653".

Figure 46 shows the amino acid sequence (SEQ ID NO:46) derived from the coding sequence of SEQ ID NO:45 shown in Figure 45.

Figure 47 shows a nucleotide sequence (SEQ ID NO:47) of a native sequence PRO4313 cDNA, wherein SEQ ID NO:47 is a clone designated herein as "DNA62312-2558".

Figure 48 shows the amino acid sequence (SEQ ID NO:48) derived from the coding sequence of SEQ ID NO:47 shown in Figure 47.

5 Figure 49 shows a nucleotide sequence (SEQ ID NO:49) of a native sequence PRO4799 cDNA, wherein SEQ ID NO:49 is a clone designated herein as "DNA62849-1604".

Figure 50 shows the amino acid sequence (SEQ ID NO:50) derived from the coding sequence of SEQ ID NO:49 shown in Figure 49.

Figure 51 shows a nucleotide sequence (SEQ ID NO:51) of a native sequence PRO4995 cDNA, wherein SEQ ID NO:51 is a clone designated herein as "DNA66307-2661".

10 Figure 52 shows the amino acid sequence (SEQ ID NO:52) derived from the coding sequence of SEQ ID NO:51 shown in Figure 51.

Figure 53 shows a nucleotide sequence (SEQ ID NO:53) of a native sequence PRO1341 cDNA, wherein SEQ ID NO:53 is a clone designated herein as "DNA66677-2535".

15 Figure 54 shows the amino acid sequence (SEQ ID NO:54) derived from the coding sequence of SEQ ID NO:53 shown in Figure 53.

Figure 55 shows a nucleotide sequence (SEQ ID NO:55) of a native sequence PRO1777 cDNA, wherein SEQ ID NO:55 is a clone designated herein as "DNA71235-1706".

Figure 56 shows the amino acid sequence (SEQ ID NO:56) derived from the coding sequence of SEQ ID NO:55 shown in Figure 55.

20 Figure 57 shows a nucleotide sequence (SEQ ID NO:57) of a native sequence PRO3580 cDNA, wherein SEQ ID NO:57 is a clone designated herein as "DNA71289-2547".

Figure 58 shows the amino acid sequence (SEQ ID NO:58) derived from the coding sequence of SEQ ID NO:57 shown in Figure 57.

25 Figure 59 shows a nucleotide sequence (SEQ ID NO:59) of a native sequence PRO1779 cDNA, wherein SEQ ID NO:59 is a clone designated herein as "DNA73775-1707".

Figure 60 shows the amino acid sequence (SEQ ID NO:60) derived from the coding sequence of SEQ ID NO:59 shown in Figure 59.

Figure 61 shows a nucleotide sequence (SEQ ID NO:61) of a native sequence PRO1754 cDNA, wherein SEQ ID NO:61 is a clone designated herein as "DNA76385-1692".

30 Figure 62 shows the amino acid sequence (SEQ ID NO:62) derived from the coding sequence of SEQ ID NO:61 shown in Figure 61.

Figure 63 shows a nucleotide sequence (SEQ ID NO:63) of a native sequence PRO1906 cDNA, wherein SEQ ID NO:63 is a clone designated herein as "DNA76395-2527".

35 Figure 64 shows the amino acid sequence (SEQ ID NO:64) derived from the coding sequence of SEQ ID NO:63 shown in Figure 63.

Figure 65 shows a nucleotide sequence (SEQ ID NO:65) of a native sequence PRO1870 cDNA, wherein SEQ ID NO:65 is a clone designated herein as "DNA77622-2516".

Figure 66 shows the amino acid sequence (SEQ ID NO:66) derived from the coding sequence of SEQ ID NO:65 shown in Figure 65.

Figure 67 shows a nucleotide sequence (SEQ ID NO:67) of a native sequence PRO4329 cDNA, wherein SEQ ID NO:67 is a clone designated herein as "DNA77629-2573".

5 Figure 68 shows the amino acid sequence (SEQ ID NO:68) derived from the coding sequence of SEQ ID NO:67 shown in Figure 67.

Figure 69 shows a nucleotide sequence (SEQ ID NO:69) of a native sequence PRO4979 cDNA, wherein SEQ ID NO:69 is a clone designated herein as "DNA77645-2648".

Figure 70 shows the amino acid sequence (SEQ ID NO:70) derived from the coding sequence of SEQ ID NO:69 shown in Figure 69.

10 Figure 71 shows a nucleotide sequence (SEQ ID NO:71) of a native sequence PRO1885 cDNA, wherein SEQ ID NO:71 is a clone designated herein as "DNA79302-2521".

Figure 72 shows the amino acid sequence (SEQ ID NO:72) derived from the coding sequence of SEQ ID NO:71 shown in Figure 71.

15 Figure 73 shows a nucleotide sequence (SEQ ID NO:73) of a native sequence PRO1882 cDNA, wherein SEQ ID NO:73 is a clone designated herein as "DNA79865-2519".

Figure 74 shows the amino acid sequence (SEQ ID NO:74) derived from the coding sequence of SEQ ID NO:73 shown in Figure 73.

Figure 75 shows a nucleotide sequence (SEQ ID NO:75) of a native sequence PRO4989 cDNA, wherein SEQ ID NO:75 is a clone designated herein as "DNA80135-2655".

20 Figure 76 shows the amino acid sequence (SEQ ID NO:76) derived from the coding sequence of SEQ ID NO:75 shown in Figure 75.

Figure 77 shows a nucleotide sequence (SEQ ID NO:77) of a native sequence PRO4323 cDNA, wherein SEQ ID NO:77 is a clone designated herein as "DNA80794-2568".

25 Figure 78 shows the amino acid sequence (SEQ ID NO:78) derived from the coding sequence of SEQ ID NO:77 shown in Figure 77.

Figure 79 shows a nucleotide sequence (SEQ ID NO:79) of a native sequence PRO1886 cDNA, wherein SEQ ID NO:79 is a clone designated herein as "DNA80796-2523".

Figure 80 shows the amino acid sequence (SEQ ID NO:80) derived from the coding sequence of SEQ ID NO:79 shown in Figure 79.

30 Figure 81 shows a nucleotide sequence (SEQ ID NO:81) of a native sequence PRO4395 cDNA, wherein SEQ ID NO:81 is a clone designated herein as "DNA80840-2605".

Figure 82 shows the amino acid sequence (SEQ ID NO:82) derived from the coding sequence of SEQ ID NO:81 shown in Figure 81.

35 Figure 83 shows a nucleotide sequence (SEQ ID NO:83) of a native sequence PRO1782 cDNA, wherein SEQ ID NO:83 is a clone designated herein as "DNA80899-2501".

Figure 84 shows the amino acid sequence (SEQ ID NO:84) derived from the coding sequence of SEQ ID NO:83 shown in Figure 83.

Figure 85 shows a nucleotide sequence (SEQ ID NO:85) of a native sequence PRO4338 cDNA, wherein SEQ ID NO:85 is a clone designated herein as "DNA81228-2580".

Figure 86 shows the amino acid sequence (SEQ ID NO:86) derived from the coding sequence of SEQ ID NO:85 shown in Figure 85.

5 Figure 87 shows a nucleotide sequence (SEQ ID NO:87) of a native sequence PRO4341 cDNA, wherein SEQ ID NO:87 is a clone designated herein as "DNA81761-2583".

Figure 88 shows the amino acid sequence (SEQ ID NO:88) derived from the coding sequence of SEQ ID NO:87 shown in Figure 87.

Figure 89 shows a nucleotide sequence (SEQ ID NO:89) of a native sequence PRO5990 cDNA, wherein SEQ ID NO:89 is a clone designated herein as "DNA96042-2682".

10 Figure 90 shows the amino acid sequence (SEQ ID NO:90) derived from the coding sequence of SEQ ID NO:89 shown in Figure 89.

Figure 91 shows a nucleotide sequence (SEQ ID NO:91) of a native sequence PRO3438 cDNA, wherein SEQ ID NO:91 is a clone designated herein as "DNA82364-2538".

15 Figure 92 shows the amino acid sequence (SEQ ID NO:92) derived from the coding sequence of SEQ ID NO:91 shown in Figure 91.

Figure 93 shows a nucleotide sequence (SEQ ID NO:93) of a native sequence PRO4321 cDNA, wherein SEQ ID NO:93 is a clone designated herein as "DNA82424-2566".

Figure 94 shows the amino acid sequence (SEQ ID NO:94) derived from the coding sequence of SEQ ID NO:93 shown in Figure 93.

20 Figure 95 shows a nucleotide sequence (SEQ ID NO:95) of a native sequence PRO4304 cDNA, wherein SEQ ID NO:95 is a clone designated herein as "DNA82430-2557".

Figure 96 shows the amino acid sequence (SEQ ID NO:96) derived from the coding sequence of SEQ ID NO:95 shown in Figure 95.

25 Figure 97 shows a nucleotide sequence (SEQ ID NO:97) of a native sequence PRO1801 cDNA, wherein SEQ ID NO:97 is a clone designated herein as "DNA83500-2506".

Figure 98 shows the amino acid sequence (SEQ ID NO:98) derived from the coding sequence of SEQ ID NO:97 shown in Figure 97.

Figure 99 shows a nucleotide sequence (SEQ ID NO:99) of a native sequence PRO4403 cDNA, wherein SEQ ID NO:99 is a clone designated herein as "DNA83509-2612".

30 Figure 100 shows the amino acid sequence (SEQ ID NO:100) derived from the coding sequence of SEQ ID NO:99 shown in Figure 99.

Figure 101 shows a nucleotide sequence (SEQ ID NO:101) of a native sequence PRO4324 cDNA, wherein SEQ ID NO:101 is a clone designated herein as "DNA83560-2569".

35 Figure 102 shows the amino acid sequence (SEQ ID NO:102) derived from the coding sequence of SEQ ID NO:101 shown in Figure 101.

Figure 103 shows a nucleotide sequence (SEQ ID NO:103) of a native sequence PRO4303 cDNA, wherein SEQ ID NO:103 is a clone designated herein as "DNA84139-2555".

Figure 104 shows the amino acid sequence (SEQ ID NO:104) derived from the coding sequence of SEQ ID NO:103 shown in Figure 103.

Figure 105 shows a nucleotide sequence (SEQ ID NO:105) of a native sequence PRO4305 cDNA, wherein SEQ ID NO:105 is a clone designated herein as "DNA84141-2556".

5 Figure 106 shows the amino acid sequence (SEQ ID NO:106) derived from the coding sequence of SEQ ID NO:105 shown in Figure 105.

Figure 107 shows a nucleotide sequence (SEQ ID NO:107) of a native sequence PRO4404 cDNA, wherein SEQ ID NO:107 is a clone designated herein as "DNA84142-2613".

Figure 108 shows the amino acid sequence (SEQ ID NO:108) derived from the coding sequence of SEQ ID NO:107 shown in Figure 107.

10 Figure 109 shows a nucleotide sequence (SEQ ID NO:109) of a native sequence PRO1884 cDNA, wherein SEQ ID NO:109 is a clone designated herein as "DNA84318-2520".

Figure 110 shows the amino acid sequence (SEQ ID NO:110) derived from the coding sequence of SEQ ID NO:109 shown in Figure 109.

15 Figure 111 shows a nucleotide sequence (SEQ ID NO:111) of a native sequence PRO4349 cDNA, wherein SEQ ID NO:111 is a clone designated herein as "DNA84909-2590".

Figure 112 shows the amino acid sequence (SEQ ID NO:112) derived from the coding sequence of SEQ ID NO:111 shown in Figure 111.

Figure 113 shows a nucleotide sequence (SEQ ID NO:113) of a native sequence PRO4401 cDNA, wherein SEQ ID NO:113 is a clone designated herein as "DNA84912-2610".

20 Figure 114 shows the amino acid sequence (SEQ ID NO:114) derived from the coding sequence of SEQ ID NO:113 shown in Figure 113.

Figure 115 shows a nucleotide sequence (SEQ ID NO:115) of a native sequence PRO1867 cDNA, wherein SEQ ID NO:115 is a clone designated herein as "DNA84925-2514".

25 Figure 116 shows the amino acid sequence (SEQ ID NO:116) derived from the coding sequence of SEQ ID NO:115 shown in Figure 115.

Figure 117 shows a nucleotide sequence (SEQ ID NO:117) of a native sequence PRO4319 cDNA, wherein SEQ ID NO:117 is a clone designated herein as "DNA84928-2564".

Figure 118 shows the amino acid sequence (SEQ ID NO:118) derived from the coding sequence of SEQ ID NO:117 shown in Figure 117.

30 Figure 119 shows a nucleotide sequence (SEQ ID NO:119) of a native sequence PRO4991 cDNA, wherein SEQ ID NO:119 is a clone designated herein as "DNA84932-2657".

Figure 120 shows the amino acid sequence (SEQ ID NO:120) derived from the coding sequence of SEQ ID NO:119 shown in Figure 119.

35 Figure 121 shows a nucleotide sequence (SEQ ID NO:121) of a native sequence PRO4398 cDNA, wherein SEQ ID NO:121 is a clone designated herein as "DNA86592-2607".

Figure 122 shows the amino acid sequence (SEQ ID NO:122) derived from the coding sequence of SEQ ID NO:121 shown in Figure 121.

Figure 123 shows a nucleotide sequence (SEQ ID NO:123) of a native sequence PRO4346 cDNA, wherein SEQ ID NO:123 is a clone designated herein as "DNA86594-2587".

Figure 124 shows the amino acid sequence (SEQ ID NO:124) derived from the coding sequence of SEQ ID NO:123 shown in Figure 123.

5 Figure 125 shows a nucleotide sequence (SEQ ID NO:125) of a native sequence PRO4350 cDNA, wherein SEQ ID NO:125 is a clone designated herein as "DNA86647-2591".

Figure 126 shows the amino acid sequence (SEQ ID NO:126) derived from the coding sequence of SEQ ID NO:125 shown in Figure 125.

Figure 127 shows a nucleotide sequence (SEQ ID NO:127) of a native sequence PRO4318 cDNA, wherein SEQ ID NO:127 is a clone designated herein as "DNA87185-2563".

10 Figure 128 shows the amino acid sequence (SEQ ID NO:128) derived from the coding sequence of SEQ ID NO:127 shown in Figure 127.

Figure 129 shows a nucleotide sequence (SEQ ID NO:129) of a native sequence PRO4340 cDNA, wherein SEQ ID NO:129 is a clone designated herein as "DNA87656-2582".

15 Figure 130 shows the amino acid sequence (SEQ ID NO:130) derived from the coding sequence of SEQ ID NO:129 shown in Figure 129.

Figure 131 shows a nucleotide sequence (SEQ ID NO:131) of a native sequence PRO4400 cDNA, wherein SEQ ID NO:131 is a clone designated herein as "DNA87974-2609".

Figure 132 shows the amino acid sequence (SEQ ID NO:132) derived from the coding sequence of SEQ ID NO:131 shown in Figure 131.

20 Figure 133 shows a nucleotide sequence (SEQ ID NO:133) of a native sequence PRO4320 cDNA, wherein SEQ ID NO:133 is a clone designated herein as "DNA88001-2565".

Figure 134 shows the amino acid sequence (SEQ ID NO:134) derived from the coding sequence of SEQ ID NO:133 shown in Figure 133.

25 Figure 135 shows a nucleotide sequence (SEQ ID NO:135) of a native sequence PRO4409 cDNA, wherein SEQ ID NO:135 is a clone designated herein as "DNA88004-2575".

Figure 136 shows the amino acid sequence (SEQ ID NO:136) derived from the coding sequence of SEQ ID NO:135 shown in Figure 135.

Figure 137 shows a nucleotide sequence (SEQ ID NO:137) of a native sequence PRO4399 cDNA, wherein SEQ ID NO:137 is a clone designated herein as "DNA89220-2608".

30 Figure 138 shows the amino acid sequence (SEQ ID NO:138) derived from the coding sequence of SEQ ID NO:137 shown in Figure 137.

Figure 139 shows a nucleotide sequence (SEQ ID NO:139) of a native sequence PRO4418 cDNA, wherein SEQ ID NO:139 is a clone designated herein as "DNA89947-2618".

35 Figure 140 shows the amino acid sequence (SEQ ID NO:140) derived from the coding sequence of SEQ ID NO:139 shown in Figure 139.

Figure 141 shows a nucleotide sequence (SEQ ID NO:141) of a native sequence PRO4330 cDNA, wherein SEQ ID NO:141 is a clone designated herein as "DNA90842-2574".

Figure 142 shows the amino acid sequence (SEQ ID NO:142) derived from the coding sequence of SEQ ID NO:141 shown in Figure 141.

Figure 143 shows a nucleotide sequence (SEQ ID NO:143) of a native sequence PRO4339 cDNA, wherein SEQ ID NO:143 is a clone designated herein as "DNA91775-2581".

5 Figure 144 shows the amino acid sequence (SEQ ID NO:144) derived from the coding sequence of SEQ ID NO:143 shown in Figure 143.

Figure 145 shows a nucleotide sequence (SEQ ID NO:145) of a native sequence PRO4326 cDNA, wherein SEQ ID NO:145 is a clone designated herein as "DNA91779-2571".

Figure 146 shows the amino acid sequence (SEQ ID NO:146) derived from the coding sequence of SEQ ID NO:145 shown in Figure 145.

10 Figure 147 shows a nucleotide sequence (SEQ ID NO:147) of a native sequence PRO6014 cDNA, wherein SEQ ID NO:147 is a clone designated herein as "DNA92217-2697".

Figure 148 shows the amino acid sequence (SEQ ID NO:148) derived from the coding sequence of SEQ ID NO:147 shown in Figure 147.

15 Figure 149 shows a nucleotide sequence (SEQ ID NO:149) of a native sequence PRO3446 cDNA, wherein SEQ ID NO:149 is a clone designated herein as "DNA92219-2541".

Figure 150 shows the amino acid sequence (SEQ ID NO:150) derived from the coding sequence of SEQ ID NO:149 shown in Figure 149.

Figure 151 shows a nucleotide sequence (SEQ ID NO:151) of a native sequence PRO4322 cDNA, wherein SEQ ID NO:151 is a clone designated herein as "DNA92223-2567".

20 Figure 152 shows the amino acid sequence (SEQ ID NO:152) derived from the coding sequence of SEQ ID NO:151 shown in Figure 151.

Figure 153 shows a nucleotide sequence (SEQ ID NO:153) of a native sequence PRO4381 cDNA, wherein SEQ ID NO:153 is a clone designated herein as "DNA92225-2603".

25 Figure 154 shows the amino acid sequence (SEQ ID NO:154) derived from the coding sequence of SEQ ID NO:153 shown in Figure 153.

Figure 155 shows a nucleotide sequence (SEQ ID NO:155) of a native sequence PRO4348 cDNA, wherein SEQ ID NO:155 is a clone designated herein as "DNA92232-2589".

Figure 156 shows the amino acid sequence (SEQ ID NO:156) derived from the coding sequence of SEQ ID NO:155 shown in Figure 155.

30 Figure 157 shows a nucleotide sequence (SEQ ID NO:157) of a native sequence PRO4371 cDNA, wherein SEQ ID NO:157 is a clone designated herein as "DNA92233-2599".

Figure 158 shows the amino acid sequence (SEQ ID NO:158) derived from the coding sequence of SEQ ID NO:157 shown in Figure 157.

35 Figure 159 shows a nucleotide sequence (SEQ ID NO:159) of a native sequence PRO3742 cDNA, wherein SEQ ID NO:159 is a clone designated herein as "DNA92243-2549".

Figure 160 shows the amino acid sequence (SEQ ID NO:160) derived from the coding sequence of SEQ ID NO:159 shown in Figure 159.

Figure 161 shows a nucleotide sequence (SEQ ID NO:161) of a native sequence PRO5773 cDNA, wherein SEQ ID NO:161 is a clone designated herein as "DNA92253-2671".

Figure 162 shows the amino acid sequence (SEQ ID NO:162) derived from the coding sequence of SEQ ID NO:161 shown in Figure 161.

5 Figure 163 shows a nucleotide sequence (SEQ ID NO:163) of a native sequence PRO5774 cDNA, wherein SEQ ID NO:163 is a clone designated herein as "DNA92254-2672".

Figure 164 shows the amino acid sequence (SEQ ID NO:164) derived from the coding sequence of SEQ ID NO:163 shown in Figure 163.

Figure 165 shows a nucleotide sequence (SEQ ID NO:165) of a native sequence PRO4343 cDNA, wherein SEQ ID NO:165 is a clone designated herein as "DNA92255-2584".

10 Figure 166 shows the amino acid sequence (SEQ ID NO:166) derived from the coding sequence of SEQ ID NO:165 shown in Figure 165.

Figure 167 shows a nucleotide sequence (SEQ ID NO:167) of a native sequence PRO4325 cDNA, wherein SEQ ID NO:167 is a clone designated herein as "DNA92269-2570".

15 Figure 168 shows the amino acid sequence (SEQ ID NO:168) derived from the coding sequence of SEQ ID NO:167 shown in Figure 167.

Figure 169 shows a nucleotide sequence (SEQ ID NO:169) of a native sequence PRO4347 cDNA, wherein SEQ ID NO:169 is a clone designated herein as "DNA92288-2588".

Figure 170 shows the amino acid sequence (SEQ ID NO:170) derived from the coding sequence of SEQ ID NO:169 shown in Figure 169.

20 Figure 171 shows a nucleotide sequence (SEQ ID NO:171) of a native sequence PRO3743 cDNA, wherein SEQ ID NO:171 is a clone designated herein as "DNA92290-2550".

Figure 172 shows the amino acid sequence (SEQ ID NO:172) derived from the coding sequence of SEQ ID NO:171 shown in Figure 171.

25 Figure 173 shows a nucleotide sequence (SEQ ID NO:173) of a native sequence PRO4426 cDNA, wherein SEQ ID NO:173 is a clone designated herein as "DNA93012-2622".

Figure 174 shows the amino acid sequence (SEQ ID NO:174) derived from the coding sequence of SEQ ID NO:173 shown in Figure 173.

Figure 175 shows a nucleotide sequence (SEQ ID NO:175) of a native sequence PRO4500 cDNA, wherein SEQ ID NO:175 is a clone designated herein as "DNA93020-2642".

30 Figure 176 shows the amino acid sequence (SEQ ID NO:176) derived from the coding sequence of SEQ ID NO:175 shown in Figure 175.

Figure 177 shows a nucleotide sequence (SEQ ID NO:177) of a native sequence PRO4389 cDNA, wherein SEQ ID NO:177 is a clone designated herein as "DNA94830-2604".

35 Figure 178 shows the amino acid sequence (SEQ ID NO:178) derived from the coding sequence of SEQ ID NO:177 shown in Figure 177.

Figure 179 shows a nucleotide sequence (SEQ ID NO:179) of a native sequence PRO4337 cDNA, wherein SEQ ID NO:179 is a clone designated herein as "DNA94833-2579".

Figure 180 shows the amino acid sequence (SEQ ID NO:180) derived from the coding sequence of SEQ ID NO:179 shown in Figure 179.

Figure 181 shows a nucleotide sequence (SEQ ID NO:181) of a native sequence PRO4992 cDNA, wherein SEQ ID NO:181 is a clone designated herein as "DNA94838-2658".

Figure 182 shows the amino acid sequence (SEQ ID NO:182) derived from the coding sequence of SEQ ID NO:181 shown in Figure 181.

Figure 183 shows a nucleotide sequence (SEQ ID NO:183) of a native sequence PRO5996 cDNA, wherein SEQ ID NO:183 is a clone designated herein as "DNA94844-2686".

Figure 184 shows the amino acid sequence (SEQ ID NO:184) derived from the coding sequence of SEQ ID NO:183 shown in Figure 183.

Figure 185 shows a nucleotide sequence (SEQ ID NO:185) of a native sequence PRO4345 cDNA, wherein SEQ ID NO:185 is a clone designated herein as "DNA94854-2586".

Figure 186 shows the amino acid sequence (SEQ ID NO:186) derived from the coding sequence of SEQ ID NO:185 shown in Figure 185.

Figure 187 shows a nucleotide sequence (SEQ ID NO:187) of a native sequence PRO4978 cDNA, wherein SEQ ID NO:187 is a clone designated herein as "DNA95930".

Figure 188 shows the amino acid sequence (SEQ ID NO:188) derived from the coding sequence of SEQ ID NO:187 shown in Figure 187.

Figure 189 shows a nucleotide sequence (SEQ ID NO:189) of a native sequence PRO5780 cDNA, wherein SEQ ID NO:189 is a clone designated herein as "DNA96868-2677".

Figure 190 shows the amino acid sequence (SEQ ID NO:190) derived from the coding sequence of SEQ ID NO:189 shown in Figure 189.

Figure 191 shows a nucleotide sequence (SEQ ID NO:191) of a native sequence PRO5992 cDNA, wherein SEQ ID NO:191 is a clone designated herein as "DNA96871-2683".

Figure 192 shows the amino acid sequence (SEQ ID NO:192) derived from the coding sequence of SEQ ID NO:191 shown in Figure 191.

Figure 193 shows a nucleotide sequence (SEQ ID NO:193) of a native sequence PRO4428 cDNA, wherein SEQ ID NO:193 is a clone designated herein as "DNA96880-2624".

Figure 194 shows the amino acid sequence (SEQ ID NO:194) derived from the coding sequence of SEQ ID NO:193 shown in Figure 193.

Figure 195 shows a nucleotide sequence (SEQ ID NO:195) of a native sequence PRO4994 cDNA, wherein SEQ ID NO:195 is a clone designated herein as "DNA96986-2660".

Figure 196 shows the amino acid sequence (SEQ ID NO:196) derived from the coding sequence of SEQ ID NO:195 shown in Figure 195.

Figure 197 shows a nucleotide sequence (SEQ ID NO:197) of a native sequence PRO5995 cDNA, wherein SEQ ID NO:197 is a clone designated herein as "DNA96988-2685".

Figure 198 shows the amino acid sequence (SEQ ID NO:198) derived from the coding sequence of SEQ ID NO:197 shown in Figure 197.

Figure 199 shows a nucleotide sequence (SEQ ID NO:199) of a native sequence PRO6094 cDNA, wherein SEQ ID NO:199 is a clone designated herein as "DNA96995-2709".

Figure 200 shows the amino acid sequence (SEQ ID NO:200) derived from the coding sequence of SEQ ID NO:199 shown in Figure 199.

5 Figure 201 shows a nucleotide sequence (SEQ ID NO:201) of a native sequence PRO4317 cDNA, wherein SEQ ID NO:201 is a clone designated herein as "DNA97004-2562".

Figure 202 shows the amino acid sequence (SEQ ID NO:202) derived from the coding sequence of SEQ ID NO:201 shown in Figure 201.

Figure 203 shows a nucleotide sequence (SEQ ID NO:203) of a native sequence PRO5997 cDNA, wherein SEQ ID NO:203 is a clone designated herein as "DNA97005-2687".

10 Figure 204 shows the amino acid sequence (SEQ ID NO:204) derived from the coding sequence of SEQ ID NO:203 shown in Figure 203.

Figure 205 shows a nucleotide sequence (SEQ ID NO:205) of a native sequence PRO5005 cDNA, wherein SEQ ID NO:205 is a clone designated herein as "DNA97009-2668".

15 Figure 206 shows the amino acid sequence (SEQ ID NO:206) derived from the coding sequence of SEQ ID NO:205 shown in Figure 205.

Figure 207 shows a nucleotide sequence (SEQ ID NO:207) of a native sequence PRO5004 cDNA, wherein SEQ ID NO:207 is a clone designated herein as "DNA97013-2667".

Figure 208 shows the amino acid sequence (SEQ ID NO:208) derived from the coding sequence of SEQ ID NO:207 shown in Figure 207.

20 Figure 209 shows a nucleotide sequence (SEQ ID NO:209) of a native sequence PRO6001 cDNA, wherein SEQ ID NO:209 is a clone designated herein as "DNA98380-2690".

Figure 210 shows the amino acid sequence (SEQ ID NO:210) derived from the coding sequence of SEQ ID NO:209 shown in Figure 209.

25 Figure 211 shows a nucleotide sequence (SEQ ID NO:211) of a native sequence PRO6013 cDNA, wherein SEQ ID NO:211 is a clone designated herein as "DNA98561-2696".

Figure 212 shows the amino acid sequence (SEQ ID NO:212) derived from the coding sequence of SEQ ID NO:211 shown in Figure 211.

Figure 213 shows a nucleotide sequence (SEQ ID NO:213) of a native sequence PRO4502 cDNA, wherein SEQ ID NO:213 is a clone designated herein as "DNA98575-2644".

30 Figure 214 shows the amino acid sequence (SEQ ID NO:214) derived from the coding sequence of SEQ ID NO:213 shown in Figure 213.

Figure 215 shows a nucleotide sequence (SEQ ID NO:215) of a native sequence PRO6007 cDNA, wherein SEQ ID NO:215 is a clone designated herein as "DNA98593-2694".

35 Figure 216 shows the amino acid sequence (SEQ ID NO:216) derived from the coding sequence of SEQ ID NO:215 shown in Figure 215.

Figure 217 shows a nucleotide sequence (SEQ ID NO:217) of a native sequence PRO6028 cDNA, wherein SEQ ID NO:217 is a clone designated herein as "DNA98600-2703".

Figure 218 shows the amino acid sequence (SEQ ID NO:218) derived from the coding sequence of SEQ ID NO:217 shown in Figure 217.

Figure 219 shows a nucleotide sequence (SEQ ID NO:219) of a native sequence PRO100 cDNA, wherein SEQ ID NO:219 is a clone designated herein as "DNA99333".

5 Figure 220 shows the amino acid sequence (SEQ ID NO:220) derived from the coding sequence of SEQ ID NO:219 shown in Figure 219.

Figure 221 shows a nucleotide sequence (SEQ ID NO:221) of a native sequence PRO4327 cDNA, wherein SEQ ID NO:221 is a clone designated herein as "DNA99391-2572".

Figure 222 shows the amino acid sequence (SEQ ID NO:222) derived from the coding sequence of SEQ ID NO:221 shown in Figure 221.

10 Figure 223 shows a nucleotide sequence (SEQ ID NO:223) of a native sequence PRO4315 cDNA, wherein SEQ ID NO:223 is a clone designated herein as "DNA99393-2560".

Figure 224 shows the amino acid sequence (SEQ ID NO:224) derived from the coding sequence of SEQ ID NO:223 shown in Figure 223.

15 Figure 225 shows a nucleotide sequence (SEQ ID NO:225) of a native sequence PRO5993 cDNA, wherein SEQ ID NO:225 is a clone designated herein as "DNA100276-2684".

Figure 226 shows the amino acid sequence (SEQ ID NO:226) derived from the coding sequence of SEQ ID NO:225 shown in Figure 225.

Figure 227 shows a nucleotide sequence (SEQ ID NO:227) of a native sequence PRO4503 cDNA, wherein SEQ ID NO:227 is a clone designated herein as "DNA100312-2645".

20 Figure 228 shows the amino acid sequence (SEQ ID NO:228) derived from the coding sequence of SEQ ID NO:227 shown in Figure 227.

Figure 229 shows a nucleotide sequence (SEQ ID NO:229) of a native sequence PRO4976 cDNA, wherein SEQ ID NO:229 is a clone designated herein as "DNA100902-2646".

25 Figure 230 shows the amino acid sequence (SEQ ID NO:230) derived from the coding sequence of SEQ ID NO:229 shown in Figure 229.

Figure 231 shows a nucleotide sequence (SEQ ID NO:231) of a native sequence PRO5798 cDNA, wherein SEQ ID NO:231 is a clone designated herein as "DNA102899-2679".

Figure 232 shows the amino acid sequence (SEQ ID NO:232) derived from the coding sequence of SEQ ID NO:231 shown in Figure 231.

30 Figure 233 shows a nucleotide sequence (SEQ ID NO:233) of a native sequence PRO6242 cDNA, wherein SEQ ID NO:233 is a clone designated herein as "DNA104875-2720".

Figure 234 shows the amino acid sequence (SEQ ID NO:234) derived from the coding sequence of SEQ ID NO:233 shown in Figure 233.

35 Figure 235 shows a nucleotide sequence (SEQ ID NO:235) of a native sequence PRO6095 cDNA, wherein SEQ ID NO:235 is a clone designated herein as "DNA105680-2710".

Figure 236 shows the amino acid sequence (SEQ ID NO:236) derived from the coding sequence of SEQ ID NO:235 shown in Figure 235.

Figure 237 shows a nucleotide sequence (SEQ ID NO:237) of a native sequence PRO6093 cDNA, wherein SEQ ID NO:237 is a clone designated herein as "DNA105779-2708".

Figure 238 shows the amino acid sequence (SEQ ID NO:238) derived from the coding sequence of SEQ ID NO:237 shown in Figure 237.

5 Figure 239 shows a nucleotide sequence (SEQ ID NO:239) of a native sequence PRO6012 cDNA, wherein SEQ ID NO:239 is a clone designated herein as "DNA105794-2695".

Figure 240 shows the amino acid sequence (SEQ ID NO:240) derived from the coding sequence of SEQ ID NO:239 shown in Figure 239.

Figure 241 shows a nucleotide sequence (SEQ ID NO:241) of a native sequence PRO6027 cDNA, wherein SEQ ID NO:241 is a clone designated herein as "DNA105838-2702".

10 Figure 242 shows the amino acid sequence (SEQ ID NO:242) derived from the coding sequence of SEQ ID NO:241 shown in Figure 241.

Figure 243 shows a nucleotide sequence (SEQ ID NO:243) of a native sequence PRO6181 cDNA, wherein SEQ ID NO:243 is a clone designated herein as "DNA107698-2715".

15 Figure 244 shows the amino acid sequence (SEQ ID NO:244) derived from the coding sequence of SEQ ID NO:243 shown in Figure 243.

Figure 245 shows a nucleotide sequence (SEQ ID NO:245) of a native sequence PRO6097 cDNA, wherein SEQ ID NO:245 is a clone designated herein as "DNA107701-2711".

Figure 246 shows the amino acid sequence (SEQ ID NO:246) derived from the coding sequence of SEQ ID NO:245 shown in Figure 245.

20 Figure 247 shows a nucleotide sequence (SEQ ID NO:247) of a native sequence PRO6090 cDNA, wherein SEQ ID NO:247 is a clone designated herein as "DNA107781-2707".

Figure 248 shows the amino acid sequence (SEQ ID NO:248) derived from the coding sequence of SEQ ID NO:247 shown in Figure 247.

25 Figure 249 shows a nucleotide sequence (SEQ ID NO:249) of a native sequence PRO7171 cDNA, wherein SEQ ID NO:249 is a clone designated herein as "DNA108670-2744".

Figure 250 shows the amino acid sequence (SEQ ID NO:250) derived from the coding sequence of SEQ ID NO:249 shown in Figure 249.

Figure 251 shows a nucleotide sequence (SEQ ID NO:251) of a native sequence PRO6258 cDNA, wherein SEQ ID NO:251 is a clone designated herein as "DNA108688-2725".

30 Figure 252 shows the amino acid sequence (SEQ ID NO:252) derived from the coding sequence of SEQ ID NO:251 shown in Figure 251.

Figure 253 shows a nucleotide sequence (SEQ ID NO:253) of a native sequence PRO9820 cDNA, wherein SEQ ID NO:253 is a clone designated herein as "DNA108769-2765".

35 Figure 254 shows the amino acid sequence (SEQ ID NO:254) derived from the coding sequence of SEQ ID NO:253 shown in Figure 253.

Figure 255 shows a nucleotide sequence (SEQ ID NO:255) of a native sequence PRO6243 cDNA, wherein SEQ ID NO:255 is a clone designated herein as "DNA108935-2721".

Figure 256 shows the amino acid sequence (SEQ ID NO:256) derived from the coding sequence of SEQ ID NO:255 shown in Figure 255.

Figure 257 shows a nucleotide sequence (SEQ ID NO:257) of a native sequence PRO6182 cDNA, wherein SEQ ID NO:257 is a clone designated herein as "DNA110700-2716".

5 Figure 258 shows the amino acid sequence (SEQ ID NO:258) derived from the coding sequence of SEQ ID NO:257 shown in Figure 257.

Figure 259 shows a nucleotide sequence (SEQ ID NO:259) of a native sequence PRO6079 cDNA, wherein SEQ ID NO:259 is a clone designated herein as "DNA111750-2706".

Figure 260 shows the amino acid sequence (SEQ ID NO:260) derived from the coding sequence of SEQ ID NO:259 shown in Figure 259.

10 Figure 261 shows a nucleotide sequence (SEQ ID NO:261) of a native sequence PRO7434 cDNA, wherein SEQ ID NO:261 is a clone designated herein as "DNA123430-2755".

Figure 262 shows the amino acid sequence (SEQ ID NO:262) derived from the coding sequence of SEQ ID NO:261 shown in Figure 261.

15 Figure 263 shows a nucleotide sequence (SEQ ID NO:263) of a native sequence PRO9865 cDNA, wherein SEQ ID NO:263 is a clone designated herein as "DNA125154-2785".

Figure 264 shows the amino acid sequence (SEQ ID NO:264) derived from the coding sequence of SEQ ID NO:263 shown in Figure 263.

Figure 265 shows a nucleotide sequence (SEQ ID NO:265) of a native sequence PRO9828 cDNA, wherein SEQ ID NO:265 is a clone designated herein as "DNA142238-2768".

20 Figure 266 shows the amino acid sequence (SEQ ID NO:266) derived from the coding sequence of SEQ ID NO:265 shown in Figure 265.

Figure 267 shows a nucleotide sequence (SEQ ID NO:267) of a native sequence PRO196 cDNA, wherein SEQ ID NO:267 is a clone designated herein as "DNA22779-1130".

25 Figure 268 shows the amino acid sequence (SEQ ID NO:268) derived from the coding sequence of SEQ ID NO:267 shown in Figure 267.

Figure 269 shows a nucleotide sequence (SEQ ID NO:269) of a native sequence PRO197 cDNA, wherein SEQ ID NO:269 is a clone designated herein as "DNA22780-1078".

Figure 270 shows the amino acid sequence (SEQ ID NO:270) derived from the coding sequence of SEQ ID NO:269 shown in Figure 269.

30 Figure 271 shows a nucleotide sequence (SEQ ID NO:271) of a native sequence PRO195 cDNA, wherein SEQ ID NO:271 is a clone designated herein as "DNA26847-1395".

Figure 272 shows the amino acid sequence (SEQ ID NO:272) derived from the coding sequence of SEQ ID NO:271 shown in Figure 271.

35 Figure 273 shows a nucleotide sequence (SEQ ID NO:273) of a native sequence PRO187 cDNA, wherein SEQ ID NO:273 is a clone designated herein as "DNA27864-1155".

Figure 274 shows the amino acid sequence (SEQ ID NO:274) derived from the coding sequence of SEQ ID NO:273 shown in Figure 273.

Figure 275 shows a nucleotide sequence (SEQ ID NO:275) of a native sequence PRO182 cDNA, wherein SEQ ID NO:275 is a clone designated herein as "DNA27865-1091".

Figure 276 shows the amino acid sequence (SEQ ID NO:276) derived from the coding sequence of SEQ ID NO:275 shown in Figure 275.

5 Figure 277 shows a nucleotide sequence (SEQ ID NO:277) of a native sequence PRO188 cDNA, wherein SEQ ID NO:277 is a clone designated herein as "DNA28497-1130".

Figure 278 shows the amino acid sequence (SEQ ID NO:278) derived from the coding sequence of SEQ ID NO:277 shown in Figure 277.

Figure 279 shows a nucleotide sequence (SEQ ID NO:279) of a native sequence PRO183 cDNA, wherein SEQ ID NO:279 is a clone designated herein as "DNA28498".

10 Figure 280 shows the amino acid sequence (SEQ ID NO:280) derived from the coding sequence of SEQ ID NO:279 shown in Figure 279.

Figure 281 shows a nucleotide sequence (SEQ ID NO:281) of a native sequence PRO184 cDNA, wherein SEQ ID NO:281 is a clone designated herein as "DNA28500".

15 Figure 282 shows the amino acid sequence (SEQ ID NO:282) derived from the coding sequence of SEQ ID NO:281 shown in Figure 281.

Figure 283 shows a nucleotide sequence (SEQ ID NO:283) of a native sequence PRO185 cDNA, wherein SEQ ID NO:283 is a clone designated herein as "DNA28503".

Figure 284 shows the amino acid sequence (SEQ ID NO:284) derived from the coding sequence of SEQ ID NO:283 shown in Figure 283.

20 Figure 285 shows a nucleotide sequence (SEQ ID NO:285) of a native sequence PRO200 cDNA, wherein SEQ ID NO:285 is a clone designated herein as "DNA29101-1122".

Figure 286 shows the amino acid sequence (SEQ ID NO:286) derived from the coding sequence of SEQ ID NO:285 shown in Figure 285.

25 Figure 287 shows a nucleotide sequence (SEQ ID NO:287) of a native sequence PRO202 cDNA, wherein SEQ ID NO:287 is a clone designated herein as "DNA30869".

Figure 288 shows the amino acid sequence (SEQ ID NO:288) derived from the coding sequence of SEQ ID NO:287 shown in Figure 287.

Figure 289 shows a nucleotide sequence (SEQ ID NO:289) of a native sequence PRO214 cDNA, wherein SEQ ID NO:289 is a clone designated herein as "DNA32286-1191".

30 Figure 290 shows the amino acid sequence (SEQ ID NO:290) derived from the coding sequence of SEQ ID NO:289 shown in Figure 289.

Figure 291 shows a nucleotide sequence (SEQ ID NO:291) of a native sequence PRO215 cDNA, wherein SEQ ID NO:291 is a clone designated herein as "DNA32288-1132".

35 Figure 292 shows the amino acid sequence (SEQ ID NO:292) derived from the coding sequence of SEQ ID NO:291 shown in Figure 291.

Figure 293 shows a nucleotide sequence (SEQ ID NO:293) of a native sequence PRO219 cDNA, wherein SEQ ID NO:293 is a clone designated herein as "DNA32290-1164".

Figure 294 shows the amino acid sequence (SEQ ID NO:294) derived from the coding sequence of SEQ ID NO:293 shown in Figure 293.

Figure 295 shows a nucleotide sequence (SEQ ID NO:295) of a native sequence PRO211 cDNA, wherein SEQ ID NO:295 is a clone designated herein as "DNA32292-1131".

Figure 296 shows the amino acid sequence (SEQ ID NO:296) derived from the coding sequence of SEQ ID NO:295 shown in Figure 295.

Figure 297 shows a nucleotide sequence (SEQ ID NO:297) of a native sequence PRO220 cDNA, wherein SEQ ID NO:297 is a clone designated herein as "DNA32298-1132".

Figure 298 shows the amino acid sequence (SEQ ID NO:298) derived from the coding sequence of SEQ ID NO:297 shown in Figure 297.

Figure 299 shows a nucleotide sequence (SEQ ID NO:299) of a native sequence PRO366 cDNA, wherein SEQ ID NO:299 is a clone designated herein as "DNA33085-1110".

Figure 300 shows the amino acid sequence (SEQ ID NO:300) derived from the coding sequence of SEQ ID NO:299 shown in Figure 299.

Figure 301 shows a nucleotide sequence (SEQ ID NO:301) of a native sequence PRO216 cDNA, wherein SEQ ID NO:301 is a clone designated herein as "DNA33087-1158".

Figure 302 shows the amino acid sequence (SEQ ID NO:302) derived from the coding sequence of SEQ ID NO:301 shown in Figure 301.

Figure 303 shows a nucleotide sequence (SEQ ID NO:303) of a native sequence PRO221 cDNA, wherein SEQ ID NO:303 is a clone designated herein as "DNA33089-1132".

Figure 304 shows the amino acid sequence (SEQ ID NO:304) derived from the coding sequence of SEQ ID NO:303 shown in Figure 303.

Figure 305 shows a nucleotide sequence (SEQ ID NO:305) of a native sequence PRO228 cDNA, wherein SEQ ID NO:305 is a clone designated herein as "DNA33092-1202".

Figure 306 shows the amino acid sequence (SEQ ID NO:306) derived from the coding sequence of SEQ ID NO:305 shown in Figure 305.

Figure 307 shows a nucleotide sequence (SEQ ID NO:307) of a native sequence PRO217 cDNA, wherein SEQ ID NO:307 is a clone designated herein as "DNA33094-1131".

Figure 308 shows the amino acid sequence (SEQ ID NO:308) derived from the coding sequence of SEQ ID NO:307 shown in Figure 307.

Figure 309 shows a nucleotide sequence (SEQ ID NO:309) of a native sequence PRO222 cDNA, wherein SEQ ID NO:309 is a clone designated herein as "DNA33107-1135".

Figure 310 shows the amino acid sequence (SEQ ID NO:310) derived from the coding sequence of SEQ ID NO:309 shown in Figure 309.

Figure 311 shows a nucleotide sequence (SEQ ID NO:311) of a native sequence PRO224 cDNA, wherein SEQ ID NO:311 is a clone designated herein as "DNA33221-1133".

Figure 312 shows the amino acid sequence (SEQ ID NO:312) derived from the coding sequence of SEQ ID NO:311 shown in Figure 311.

Figure 313 shows a nucleotide sequence (SEQ ID NO:313) of a native sequence PRO230 cDNA, wherein SEQ ID NO:313 is a clone designated herein as "DNA33223-1136".

Figure 314 shows the amino acid sequence (SEQ ID NO:314) derived from the coding sequence of SEQ ID NO:313 shown in Figure 313.

5 Figure 315 shows a nucleotide sequence (SEQ ID NO:315) of a native sequence PRO198 cDNA, wherein SEQ ID NO:315 is a clone designated herein as "DNA33457-1078".

Figure 316 shows the amino acid sequence (SEQ ID NO:316) derived from the coding sequence of SEQ ID NO:315 shown in Figure 315.

Figure 317 shows a nucleotide sequence (SEQ ID NO:317) of a native sequence PRO226 cDNA, wherein SEQ ID NO:317 is a clone designated herein as "DNA33460-1166".

10 Figure 318 shows the amino acid sequence (SEQ ID NO:318) derived from the coding sequence of SEQ ID NO:317 shown in Figure 317.

Figure 319 shows a nucleotide sequence (SEQ ID NO:319) of a native sequence PRO261 cDNA, wherein SEQ ID NO:319 is a clone designated herein as "DNA33473-1176".

15 Figure 320 shows the amino acid sequence (SEQ ID NO:320) derived from the coding sequence of SEQ ID NO:319 shown in Figure 319.

Figure 321 shows a nucleotide sequence (SEQ ID NO:321) of a native sequence PRO242 cDNA, wherein SEQ ID NO:321 is a clone designated herein as "DNA33785-1143".

Figure 322 shows the amino acid sequence (SEQ ID NO:322) derived from the coding sequence of SEQ ID NO:321 shown in Figure 321.

20 Figure 323 shows a nucleotide sequence (SEQ ID NO:323) of a native sequence PRO227 cDNA, wherein SEQ ID NO:323 is a clone designated herein as "DNA33786-1132".

Figure 324 shows the amino acid sequence (SEQ ID NO:324) derived from the coding sequence of SEQ ID NO:323 shown in Figure 323.

25 Figure 325 shows a nucleotide sequence (SEQ ID NO:325) of a native sequence PRO237 cDNA, wherein SEQ ID NO:325 is a clone designated herein as "DNA34353-1428".

Figure 326 shows the amino acid sequence (SEQ ID NO:326) derived from the coding sequence of SEQ ID NO:325 shown in Figure 325.

Figure 327 shows a nucleotide sequence (SEQ ID NO:327) of a native sequence PRO241 cDNA, wherein SEQ ID NO:327 is a clone designated herein as "DNA34392-1170".

30 Figure 328 shows the amino acid sequence (SEQ ID NO:328) derived from the coding sequence of SEQ ID NO:327 shown in Figure 327.

Figure 329 shows a nucleotide sequence (SEQ ID NO:329) of a native sequence PRO231 cDNA, wherein SEQ ID NO:329 is a clone designated herein as "DNA34434-1139".

35 Figure 330 shows the amino acid sequence (SEQ ID NO:330) derived from the coding sequence of SEQ ID NO:329 shown in Figure 329.

Figure 331 shows a nucleotide sequence (SEQ ID NO:331) of a native sequence PRO235 cDNA, wherein SEQ ID NO:331 is a clone designated herein as "DNA35558-1167".

Figure 332 shows the amino acid sequence (SEQ ID NO:332) derived from the coding sequence of SEQ ID NO:331 shown in Figure 331.

Figure 333 shows a nucleotide sequence (SEQ ID NO:333) of a native sequence PRO323 cDNA, wherein SEQ ID NO:333 is a clone designated herein as "DNA35595-1228".

5 Figure 334 shows the amino acid sequence (SEQ ID NO:334) derived from the coding sequence of SEQ ID NO:333 shown in Figure 333.

Figure 335 shows a nucleotide sequence (SEQ ID NO:335) of a native sequence PRO245 cDNA, wherein SEQ ID NO:335 is a clone designated herein as "DNA35638-1216".

Figure 336 shows the amino acid sequence (SEQ ID NO:336) derived from the coding sequence of SEQ ID NO:335 shown in Figure 335.

10 Figure 337 shows a nucleotide sequence (SEQ ID NO:337) of a native sequence PRO246 cDNA, wherein SEQ ID NO:337 is a clone designated herein as "DNA35639-1172".

Figure 338 shows the amino acid sequence (SEQ ID NO:338) derived from the coding sequence of SEQ ID NO:337 shown in Figure 337.

15 Figure 339 shows a nucleotide sequence (SEQ ID NO:339) of a native sequence PRO288 cDNA, wherein SEQ ID NO:339 is a clone designated herein as "DNA35663-1129".

Figure 340 shows the amino acid sequence (SEQ ID NO:340) derived from the coding sequence of SEQ ID NO:339 shown in Figure 339.

Figure 341 shows a nucleotide sequence (SEQ ID NO:341) of a native sequence PRO248 cDNA, wherein SEQ ID NO:341 is a clone designated herein as "DNA35674-1142".

20 Figure 342 shows the amino acid sequence (SEQ ID NO:342) derived from the coding sequence of SEQ ID NO:341 shown in Figure 341.

Figure 343 shows a nucleotide sequence (SEQ ID NO:343) of a native sequence PRO257 cDNA, wherein SEQ ID NO:343 is a clone designated herein as "DNA35841-1173".

25 Figure 344 shows the amino acid sequence (SEQ ID NO:344) derived from the coding sequence of SEQ ID NO:343 shown in Figure 343.

Figure 345 shows a nucleotide sequence (SEQ ID NO:345) of a native sequence PRO172 cDNA, wherein SEQ ID NO:345 is a clone designated herein as "DNA35916-1161".

Figure 346 shows the amino acid sequence (SEQ ID NO:346) derived from the coding sequence of SEQ ID NO:345 shown in Figure 345.

30 Figure 347 shows a nucleotide sequence (SEQ ID NO:347) of a native sequence PRO258 cDNA, wherein SEQ ID NO:347 is a clone designated herein as "DNA35918-1174".

Figure 348 shows the amino acid sequence (SEQ ID NO:348) derived from the coding sequence of SEQ ID NO:347 shown in Figure 347.

35 Figure 349 shows a nucleotide sequence (SEQ ID NO:349) of a native sequence PRO265 cDNA, wherein SEQ ID NO:349 is a clone designated herein as "DNA36350-1158".

Figure 350 shows the amino acid sequence (SEQ ID NO:350) derived from the coding sequence of SEQ ID NO:349 shown in Figure 349.

Figure 351 shows a nucleotide sequence (SEQ ID NO:351) of a native sequence PRO326 cDNA, wherein SEQ ID NO:351 is a clone designated herein as "DNA37140-1234".

Figure 352 shows the amino acid sequence (SEQ ID NO:352) derived from the coding sequence of SEQ ID NO:351 shown in Figure 351.

5 Figure 353 shows a nucleotide sequence (SEQ ID NO:353) of a native sequence PRO266 cDNA, wherein SEQ ID NO:353 is a clone designated herein as "DNA37150-1178".

Figure 354 shows the amino acid sequence (SEQ ID NO:354) derived from the coding sequence of SEQ ID NO:353 shown in Figure 353.

Figure 355 shows a nucleotide sequence (SEQ ID NO:355) of a native sequence PRO269 cDNA, wherein SEQ ID NO:355 is a clone designated herein as "DNA38260-1180".

10 Figure 356 shows the amino acid sequence (SEQ ID NO:356) derived from the coding sequence of SEQ ID NO:355 shown in Figure 355.

Figure 357 shows a nucleotide sequence (SEQ ID NO:357) of a native sequence PRO285 cDNA, wherein SEQ ID NO:357 is a clone designated herein as "DNA40021-1154".

15 Figure 358 shows the amino acid sequence (SEQ ID NO:358) derived from the coding sequence of SEQ ID NO:357 shown in Figure 357.

Figure 359 shows a nucleotide sequence (SEQ ID NO:359) of a native sequence PRO328 cDNA, wherein SEQ ID NO:359 is a clone designated herein as "DNA40587-1231".

Figure 360 shows the amino acid sequence (SEQ ID NO:360) derived from the coding sequence of SEQ ID NO:359 shown in Figure 359.

20 Figure 361 shows a nucleotide sequence (SEQ ID NO:361) of a native sequence PRO344 cDNA, wherein SEQ ID NO:361 is a clone designated herein as "DNA40592-1242".

Figure 362 shows the amino acid sequence (SEQ ID NO:362) derived from the coding sequence of SEQ ID NO:361 shown in Figure 361.

25 Figure 363 shows a nucleotide sequence (SEQ ID NO:363) of a native sequence PRO272 cDNA, wherein SEQ ID NO:363 is a clone designated herein as "DNA40620-1183".

Figure 364 shows the amino acid sequence (SEQ ID NO:364) derived from the coding sequence of SEQ ID NO:363 shown in Figure 363.

Figure 365 shows a nucleotide sequence (SEQ ID NO:365) of a native sequence PRO301 cDNA, wherein SEQ ID NO:365 is a clone designated herein as "DNA40628-1216".

30 Figure 366 shows the amino acid sequence (SEQ ID NO:366) derived from the coding sequence of SEQ ID NO:365 shown in Figure 365.

Figure 367 shows a nucleotide sequence (SEQ ID NO:367) of a native sequence PRO331 cDNA, wherein SEQ ID NO:367 is a clone designated herein as "DNA40981-1234".

35 Figure 368 shows the amino acid sequence (SEQ ID NO:368) derived from the coding sequence of SEQ ID NO:367 shown in Figure 367.

Figure 369 shows a nucleotide sequence (SEQ ID NO:369) of a native sequence PRO332 cDNA, wherein SEQ ID NO:369 is a clone designated herein as "DNA40982-1235".

Figure 370 shows the amino acid sequence (SEQ ID NO:370) derived from the coding sequence of SEQ ID NO:369 shown in Figure 369.

Figure 371 shows a nucleotide sequence (SEQ ID NO:371) of a native sequence PRO353 cDNA, wherein SEQ ID NO:371 is a clone designated herein as "DNA41234-1242".

5 Figure 372 shows the amino acid sequence (SEQ ID NO:372) derived from the coding sequence of SEQ ID NO:371 shown in Figure 371.

Figure 373 shows a nucleotide sequence (SEQ ID NO:373) of a native sequence PRO310 cDNA, wherein SEQ ID NO:373 is a clone designated herein as "DNA43046-1225".

Figure 374 shows the amino acid sequence (SEQ ID NO:374) derived from the coding sequence of SEQ ID NO:373 shown in Figure 373.

10 Figure 375 shows a nucleotide sequence (SEQ ID NO:375) of a native sequence PRO337 cDNA, wherein SEQ ID NO:375 is a clone designated herein as "DNA43316-1237".

Figure 376 shows the amino acid sequence (SEQ ID NO:376) derived from the coding sequence of SEQ ID NO:375 shown in Figure 375.

15 Figure 377 shows a nucleotide sequence (SEQ ID NO:377) of a native sequence PRO346 cDNA, wherein SEQ ID NO:377 is a clone designated herein as "DNA44167-1243".

Figure 378 shows the amino acid sequence (SEQ ID NO:378) derived from the coding sequence of SEQ ID NO:377 shown in Figure 377.

Figure 379 shows a nucleotide sequence (SEQ ID NO:379) of a native sequence PRO350 cDNA, wherein SEQ ID NO:379 is a clone designated herein as "DNA44175-1314".

20 Figure 380 shows the amino acid sequence (SEQ ID NO:380) derived from the coding sequence of SEQ ID NO:379 shown in Figure 379.

Figure 381 shows a nucleotide sequence (SEQ ID NO:381) of a native sequence PRO526 cDNA, wherein SEQ ID NO:381 is a clone designated herein as "DNA44184-1319".

25 Figure 382 shows the amino acid sequence (SEQ ID NO:382) derived from the coding sequence of SEQ ID NO:381 shown in Figure 381.

Figure 383 shows a nucleotide sequence (SEQ ID NO:383) of a native sequence PRO381 cDNA, wherein SEQ ID NO:383 is a clone designated herein as "DNA44194-1317".

Figure 384 shows the amino acid sequence (SEQ ID NO:384) derived from the coding sequence of SEQ ID NO:383 shown in Figure 383.

30 Figure 385 shows a nucleotide sequence (SEQ ID NO:385) of a native sequence PRO846 cDNA, wherein SEQ ID NO:385 is a clone designated herein as "DNA44196-1353".

Figure 386 shows the amino acid sequence (SEQ ID NO:386) derived from the coding sequence of SEQ ID NO:385 shown in Figure 385.

35 Figure 387 shows a nucleotide sequence (SEQ ID NO:387) of a native sequence PRO363 cDNA, wherein SEQ ID NO:387 is a clone designated herein as "DNA45419-1252".

Figure 388 shows the amino acid sequence (SEQ ID NO:388) derived from the coding sequence of SEQ ID NO:387 shown in Figure 387.

Figure 389 shows a nucleotide sequence (SEQ ID NO:389) of a native sequence PRO365 cDNA, wherein SEQ ID NO:389 is a clone designated herein as "DNA46777-1253".

Figure 390 shows the amino acid sequence (SEQ ID NO:390) derived from the coding sequence of SEQ ID NO:389 shown in Figure 389.

5 Figure 391 shows a nucleotide sequence (SEQ ID NO:391) of a native sequence PRO1310 cDNA, wherein SEQ ID NO:391 is a clone designated herein as "DNA47394-1572".

Figure 392 shows the amino acid sequence (SEQ ID NO:392) derived from the coding sequence of SEQ ID NO:391 shown in Figure 391.

Figure 393 shows a nucleotide sequence (SEQ ID NO:393) of a native sequence PRO731 cDNA, wherein SEQ ID NO:393 is a clone designated herein as "DNA48331-1329".

10 Figure 394 shows the amino acid sequence (SEQ ID NO:394) derived from the coding sequence of SEQ ID NO:393 shown in Figure 393.

Figure 395 shows a nucleotide sequence (SEQ ID NO:395) of a native sequence PRO322 cDNA, wherein SEQ ID NO:395 is a clone designated herein as "DNA48336-1309".

15 Figure 396 shows the amino acid sequence (SEQ ID NO:396) derived from the coding sequence of SEQ ID NO:395 shown in Figure 395.

Figure 397 shows a nucleotide sequence (SEQ ID NO:397) of a native sequence PRO536 cDNA, wherein SEQ ID NO:397 is a clone designated herein as "DNA49142-1430".

Figure 398 shows the amino acid sequence (SEQ ID NO:398) derived from the coding sequence of SEQ ID NO:397 shown in Figure 397.

20 Figure 399 shows a nucleotide sequence (SEQ ID NO:399) of a native sequence PRO719 cDNA, wherein SEQ ID NO:399 is a clone designated herein as "DNA49646-1327".

Figure 400 shows the amino acid sequence (SEQ ID NO:400) derived from the coding sequence of SEQ ID NO:399 shown in Figure 399.

25 Figure 401 shows a nucleotide sequence (SEQ ID NO:401) of a native sequence PRO619 cDNA, wherein SEQ ID NO:401 is a clone designated herein as "DNA49821-1562".

Figure 402 shows the amino acid sequence (SEQ ID NO:402) derived from the coding sequence of SEQ ID NO:401 shown in Figure 401.

Figure 403 shows a nucleotide sequence (SEQ ID NO:403) of a native sequence PRO771 cDNA, wherein SEQ ID NO:403 is a clone designated herein as "DNA49829-1346".

30 Figure 404 shows the amino acid sequence (SEQ ID NO:404) derived from the coding sequence of SEQ ID NO:403 shown in Figure 403.

Figure 405 shows a nucleotide sequence (SEQ ID NO:405) of a native sequence PRO1083 cDNA, wherein SEQ ID NO:405 is a clone designated herein as "DNA50921-1458".

35 Figure 406 shows the amino acid sequence (SEQ ID NO:406) derived from the coding sequence of SEQ ID NO:405 shown in Figure 405.

Figure 407 shows a nucleotide sequence (SEQ ID NO:407) of a native sequence PRO862 cDNA, wherein SEQ ID NO:407 is a clone designated herein as "DNA52187-1354".

Figure 408 shows the amino acid sequence (SEQ ID NO:408) derived from the coding sequence of SEQ ID NO:407 shown in Figure 407.

Figure 409 shows a nucleotide sequence (SEQ ID NO:409) of a native sequence PRO733 cDNA, wherein SEQ ID NO:409 is a clone designated herein as "DNA52196-1348".

Figure 410 shows the amino acid sequence (SEQ ID NO:410) derived from the coding sequence of SEQ ID NO:409 shown in Figure 409.

Figure 411 shows a nucleotide sequence (SEQ ID NO:411) of a native sequence PRO1188 cDNA, wherein SEQ ID NO:411 is a clone designated herein as "DNA52598-1518".

Figure 412 shows the amino acid sequence (SEQ ID NO:412) derived from the coding sequence of SEQ ID NO:411 shown in Figure 411.

Figure 413 shows a nucleotide sequence (SEQ ID NO:413) of a native sequence PRO770 cDNA, wherein SEQ ID NO:413 is a clone designated herein as "DNA54228-1366".

Figure 414 shows the amino acid sequence (SEQ ID NO:414) derived from the coding sequence of SEQ ID NO:413 shown in Figure 413.

Figure 415 shows a nucleotide sequence (SEQ ID NO:415) of a native sequence PRO1080 cDNA, wherein SEQ ID NO:415 is a clone designated herein as "DNA56047-1456".

Figure 416 shows the amino acid sequence (SEQ ID NO:416) derived from the coding sequence of SEQ ID NO:415 shown in Figure 415.

Figure 417 shows a nucleotide sequence (SEQ ID NO:417) of a native sequence PRO1017 cDNA, wherein SEQ ID NO:417 is a clone designated herein as "DNA56112-1379".

Figure 418 shows the amino acid sequence (SEQ ID NO:418) derived from the coding sequence of SEQ ID NO:417 shown in Figure 417.

Figure 419 shows a nucleotide sequence (SEQ ID NO:419) of a native sequence PRO1016 cDNA, wherein SEQ ID NO:419 is a clone designated herein as "DNA56113-1378".

Figure 420 shows the amino acid sequence (SEQ ID NO:420) derived from the coding sequence of SEQ ID NO:419 shown in Figure 419.

Figure 421 shows a nucleotide sequence (SEQ ID NO:421) of a native sequence PRO792 cDNA, wherein SEQ ID NO:421 is a clone designated herein as "DNA56352-1358".

Figure 422 shows the amino acid sequence (SEQ ID NO:422) derived from the coding sequence of SEQ ID NO:421 shown in Figure 421.

Figure 423 shows a nucleotide sequence (SEQ ID NO:423) of a native sequence PRO938 cDNA, wherein SEQ ID NO:423 is a clone designated herein as "DNA56433-1406".

Figure 424 shows the amino acid sequence (SEQ ID NO:424) derived from the coding sequence of SEQ ID NO:423 shown in Figure 423.

Figure 425 shows a nucleotide sequence (SEQ ID NO:425) of a native sequence PRO1012 cDNA, wherein SEQ ID NO:425 is a clone designated herein as "DNA56439-1376".

Figure 426 shows the amino acid sequence (SEQ ID NO:426) derived from the coding sequence of SEQ ID NO:425 shown in Figure 425.

Figure 427 shows a nucleotide sequence (SEQ ID NO:427) of a native sequence PRO1008 cDNA, wherein SEQ ID NO:427 is a clone designated herein as "DNA57530-1375".

Figure 428 shows the amino acid sequence (SEQ ID NO:428) derived from the coding sequence of SEQ ID NO:427 shown in Figure 427.

5 Figure 429 shows a nucleotide sequence (SEQ ID NO:429) of a native sequence PRO1075 cDNA, wherein SEQ ID NO:429 is a clone designated herein as "DNA57689-1385".

Figure 430 shows the amino acid sequence (SEQ ID NO:430) derived from the coding sequence of SEQ ID NO:429 shown in Figure 429.

Figure 431 shows a nucleotide sequence (SEQ ID NO:431) of a native sequence PRO1007 cDNA, wherein SEQ ID NO:431 is a clone designated herein as "DNA57690-1374".

10 Figure 432 shows the amino acid sequence (SEQ ID NO:432) derived from the coding sequence of SEQ ID NO:431 shown in Figure 431.

Figure 433 shows a nucleotide sequence (SEQ ID NO:433) of a native sequence PRO1056 cDNA, wherein SEQ ID NO:433 is a clone designated herein as "DNA57693-1424".

15 Figure 434 shows the amino acid sequence (SEQ ID NO:434) derived from the coding sequence of SEQ ID NO:433 shown in Figure 433.

Figure 435 shows a nucleotide sequence (SEQ ID NO:435) of a native sequence PRO791 cDNA, wherein SEQ ID NO:435 is a clone designated herein as "DNA57838-1337".

Figure 436 shows the amino acid sequence (SEQ ID NO:436) derived from the coding sequence of SEQ ID NO:435 shown in Figure 435.

20 Figure 437 shows a nucleotide sequence (SEQ ID NO:437) of a native sequence PRO1111 cDNA, wherein SEQ ID NO:437 is a clone designated herein as "DNA58721-1475".

Figure 438 shows the amino acid sequence (SEQ ID NO:438) derived from the coding sequence of SEQ ID NO:437 shown in Figure 437.

25 Figure 439 shows a nucleotide sequence (SEQ ID NO:439) of a native sequence PRO812 cDNA, wherein SEQ ID NO:439 is a clone designated herein as "DNA59205-1421".

Figure 440 shows the amino acid sequence (SEQ ID NO:440) derived from the coding sequence of SEQ ID NO:439 shown in Figure 439.

Figure 441 shows a nucleotide sequence (SEQ ID NO:441) of a native sequence PRO1066 cDNA, wherein SEQ ID NO:441 is a clone designated herein as "DNA59215-1425".

30 Figure 442 shows the amino acid sequence (SEQ ID NO:442) derived from the coding sequence of SEQ ID NO:441 shown in Figure 441.

Figure 443 shows a nucleotide sequence (SEQ ID NO:443) of a native sequence PRO1185 cDNA, wherein SEQ ID NO:443 is a clone designated herein as "DNA59220-1514".

35 Figure 444 shows the amino acid sequence (SEQ ID NO:444) derived from the coding sequence of SEQ ID NO:443 shown in Figure 443.

Figure 445 shows a nucleotide sequence (SEQ ID NO:445) of a native sequence PRO1031 cDNA, wherein SEQ ID NO:445 is a clone designated herein as "DNA59294-1381".

Figure 446 shows the amino acid sequence (SEQ ID NO:446) derived from the coding sequence of SEQ ID NO:445 shown in Figure 445.

Figure 447 shows a nucleotide sequence (SEQ ID NO:447) of a native sequence PRO1360 cDNA, wherein SEQ ID NO:447 is a clone designated herein as "DNA59488-1603".

5 Figure 448 shows the amino acid sequence (SEQ ID NO:448) derived from the coding sequence of SEQ ID NO:447 shown in Figure 447.

Figure 449 shows a nucleotide sequence (SEQ ID NO:449) of a native sequence PRO1309 cDNA, wherein SEQ ID NO:449 is a clone designated herein as "DNA59588-1571".

Figure 450 shows the amino acid sequence (SEQ ID NO:450) derived from the coding sequence of SEQ ID NO:449 shown in Figure 449.

10 Figure 451 shows a nucleotide sequence (SEQ ID NO:451) of a native sequence PRO1107 cDNA, wherein SEQ ID NO:451 is a clone designated herein as "DNA59606-1471".

Figure 452 shows the amino acid sequence (SEQ ID NO:452) derived from the coding sequence of SEQ ID NO:451 shown in Figure 451.

15 Figure 453 shows a nucleotide sequence (SEQ ID NO:453) of a native sequence PRO836 cDNA, wherein SEQ ID NO:453 is a clone designated herein as "DNA59620-1463".

Figure 454 shows the amino acid sequence (SEQ ID NO:454) derived from the coding sequence of SEQ ID NO:453 shown in Figure 453.

Figure 455 shows a nucleotide sequence (SEQ ID NO:455) of a native sequence PRO1132 cDNA, wherein SEQ ID NO:455 is a clone designated herein as "DNA59767-1489".

20 Figure 456 shows the amino acid sequence (SEQ ID NO:456) derived from the coding sequence of SEQ ID NO:455 shown in Figure 455.

Figure 457 shows a nucleotide sequence (SEQ ID NO:457) of a native sequence PRO1131 cDNA, wherein SEQ ID NO:457 is a clone designated herein as "DNA59777-1480".

25 Figure 458 shows the amino acid sequence (SEQ ID NO:458) derived from the coding sequence of SEQ ID NO:457 shown in Figure 457.

Figure 459 shows a nucleotide sequence (SEQ ID NO:459) of a native sequence PRO1130 cDNA, wherein SEQ ID NO:459 is a clone designated herein as "DNA59814-1486".

Figure 460 shows the amino acid sequence (SEQ ID NO:460) derived from the coding sequence of SEQ ID NO:459 shown in Figure 459.

30 Figure 461 shows a nucleotide sequence (SEQ ID NO:461) of a native sequence PRO844 cDNA, wherein SEQ ID NO:461 is a clone designated herein as "DNA59839-1461".

Figure 462 shows the amino acid sequence (SEQ ID NO:462) derived from the coding sequence of SEQ ID NO:461 shown in Figure 461.

35 Figure 463 shows a nucleotide sequence (SEQ ID NO:463) of a native sequence PRO1154 cDNA, wherein SEQ ID NO:463 is a clone designated herein as "DNA59846-1503".

Figure 464 shows the amino acid sequence (SEQ ID NO:464) derived from the coding sequence of SEQ ID NO:463 shown in Figure 463.

Figure 465 shows a nucleotide sequence (SEQ ID NO:465) of a native sequence PRO1181 cDNA, wherein SEQ ID NO:465 is a clone designated herein as "DNA59847-1511".

Figure 466 shows the amino acid sequence (SEQ ID NO:466) derived from the coding sequence of SEQ ID NO:465 shown in Figure 465.

5 Figure 467 shows a nucleotide sequence (SEQ ID NO:467) of a native sequence PRO1126 cDNA, wherein SEQ ID NO:467 is a clone designated herein as "DNA60615-1483".

Figure 468 shows the amino acid sequence (SEQ ID NO:468) derived from the coding sequence of SEQ ID NO:467 shown in Figure 467.

Figure 469 shows a nucleotide sequence (SEQ ID NO:469) of a native sequence PRO1186 cDNA, wherein SEQ ID NO:469 is a clone designated herein as "DNA60621-1516".

10 Figure 470 shows the amino acid sequence (SEQ ID NO:470) derived from the coding sequence of SEQ ID NO:469 shown in Figure 469.

Figure 471 shows a nucleotide sequence (SEQ ID NO:471) of a native sequence PRO1198 cDNA, wherein SEQ ID NO:471 is a clone designated herein as "DNA60622-1525".

15 Figure 472 shows the amino acid sequence (SEQ ID NO:472) derived from the coding sequence of SEQ ID NO:471 shown in Figure 471.

Figure 473 shows a nucleotide sequence (SEQ ID NO:473) of a native sequence PRO1159 cDNA, wherein SEQ ID NO:473 is a clone designated herein as "DNA60627-1508".

Figure 474 shows the amino acid sequence (SEQ ID NO:474) derived from the coding sequence of SEQ ID NO:473 shown in Figure 473.

20 Figure 475 shows a nucleotide sequence (SEQ ID NO:475) of a native sequence PRO1265 cDNA, wherein SEQ ID NO:475 is a clone designated herein as "DNA60764-1533".

Figure 476 shows the amino acid sequence (SEQ ID NO:476) derived from the coding sequence of SEQ ID NO:475 shown in Figure 475.

25 Figure 477 shows a nucleotide sequence (SEQ ID NO:477) of a native sequence PRO1250 cDNA, wherein SEQ ID NO:477 is a clone designated herein as "DNA60775-1532".

Figure 478 shows the amino acid sequence (SEQ ID NO:478) derived from the coding sequence of SEQ ID NO:477 shown in Figure 477.

Figure 479 shows a nucleotide sequence (SEQ ID NO:479) of a native sequence PRO1475 cDNA, wherein SEQ ID NO:479 is a clone designated herein as "DNA61185-1646".

30 Figure 480 shows the amino acid sequence (SEQ ID NO:480) derived from the coding sequence of SEQ ID NO:479 shown in Figure 479.

Figure 481 shows a nucleotide sequence (SEQ ID NO:481) of a native sequence PRO1312 cDNA, wherein SEQ ID NO:481 is a clone designated herein as "DNA61873-1574".

35 Figure 482 shows the amino acid sequence (SEQ ID NO:482) derived from the coding sequence of SEQ ID NO:481 shown in Figure 481.

Figure 483 shows a nucleotide sequence (SEQ ID NO:483) of a native sequence PRO1308 cDNA, wherein SEQ ID NO:483 is a clone designated herein as "DNA62306-1570".

Figure 484 shows the amino acid sequence (SEQ ID NO:484) derived from the coding sequence of SEQ ID NO:483 shown in Figure 483.

Figure 485 shows a nucleotide sequence (SEQ ID NO:485) of a native sequence PRO1326 cDNA, wherein SEQ ID NO:485 is a clone designated herein as "DNA62808-1582".

Figure 486 shows the amino acid sequence (SEQ ID NO:486) derived from the coding sequence of SEQ ID NO:485 shown in Figure 485.

Figure 487 shows a nucleotide sequence (SEQ ID NO:487) of a native sequence PRO1192 cDNA, wherein SEQ ID NO:487 is a clone designated herein as "DNA62814-1521".

Figure 488 shows the amino acid sequence (SEQ ID NO:488) derived from the coding sequence of SEQ ID NO:487 shown in Figure 487.

Figure 489 shows a nucleotide sequence (SEQ ID NO:489) of a native sequence PRO1246 cDNA, wherein SEQ ID NO:489 is a clone designated herein as "DNA64885-1529".

Figure 490 shows the amino acid sequence (SEQ ID NO:490) derived from the coding sequence of SEQ ID NO:489 shown in Figure 489.

Figure 491 shows a nucleotide sequence (SEQ ID NO:491) of a native sequence PRO1356 cDNA, wherein SEQ ID NO:491 is a clone designated herein as "DNA64886-1601".

Figure 492 shows the amino acid sequence (SEQ ID NO:492) derived from the coding sequence of SEQ ID NO:491 shown in Figure 491.

Figure 493 shows a nucleotide sequence (SEQ ID NO:493) of a native sequence PRO1275 cDNA, wherein SEQ ID NO:493 is a clone designated herein as "DNA64888-1542".

Figure 494 shows the amino acid sequence (SEQ ID NO:494) derived from the coding sequence of SEQ ID NO:493 shown in Figure 493.

Figure 495 shows a nucleotide sequence (SEQ ID NO:495) of a native sequence PRO1274 cDNA, wherein SEQ ID NO:495 is a clone designated herein as "DNA64889-1541".

Figure 496 shows the amino acid sequence (SEQ ID NO:496) derived from the coding sequence of SEQ ID NO:495 shown in Figure 495.

Figure 497 shows a nucleotide sequence (SEQ ID NO:497) of a native sequence PRO1358 cDNA, wherein SEQ ID NO:497 is a clone designated herein as "DNA64890-1612".

Figure 498 shows the amino acid sequence (SEQ ID NO:498) derived from the coding sequence of SEQ ID NO:497 shown in Figure 497.

Figure 499 shows a nucleotide sequence (SEQ ID NO:499) of a native sequence PRO1286 cDNA, wherein SEQ ID NO:499 is a clone designated herein as "DNA64903-1553".

Figure 500 shows the amino acid sequence (SEQ ID NO:500) derived from the coding sequence of SEQ ID NO:499 shown in Figure 499.

Figure 501 shows a nucleotide sequence (SEQ ID NO:501) of a native sequence PRO1294 cDNA, wherein SEQ ID NO:501 is a clone designated herein as "DNA64905-1558".

Figure 502 shows the amino acid sequence (SEQ ID NO:502) derived from the coding sequence of SEQ ID NO:501 shown in Figure 501.

Figure 503 shows a nucleotide sequence (SEQ ID NO:503) of a native sequence PRO1273 cDNA, wherein SEQ ID NO:503 is a clone designated herein as "DNA65402-1540".

Figure 504 shows the amino acid sequence (SEQ ID NO:504) derived from the coding sequence of SEQ ID NO:503 shown in Figure 503.

5 Figure 505 shows a nucleotide sequence (SEQ ID NO:505) of a native sequence PRO1279 cDNA, wherein SEQ ID NO:505 is a clone designated herein as "DNA65405-1547".

Figure 506 shows the amino acid sequence (SEQ ID NO:506) derived from the coding sequence of SEQ ID NO:505 shown in Figure 505.

Figure 507 shows a nucleotide sequence (SEQ ID NO:507) of a native sequence PRO1195 cDNA, wherein SEQ ID NO:507 is a clone designated herein as "DNA65412-1523".

10 Figure 508 shows the amino acid sequence (SEQ ID NO:508) derived from the coding sequence of SEQ ID NO:507 shown in Figure 507.

Figure 509 shows a nucleotide sequence (SEQ ID NO:509) of a native sequence PRO1271 cDNA, wherein SEQ ID NO:509 is a clone designated herein as "DNA66309-1538".

15 Figure 510 shows the amino acid sequence (SEQ ID NO:510) derived from the coding sequence of SEQ ID NO:509 shown in Figure 509.

Figure 511 shows a nucleotide sequence (SEQ ID NO:511) of a native sequence PRO1338 cDNA, wherein SEQ ID NO:511 is a clone designated herein as "DNA66667-1596".

Figure 512 shows the amino acid sequence (SEQ ID NO:512) derived from the coding sequence of SEQ ID NO:511 shown in Figure 511.

20 Figure 513 shows a nucleotide sequence (SEQ ID NO:513) of a native sequence PRO1343 cDNA, wherein SEQ ID NO:513 is a clone designated herein as "DNA66675-1587".

Figure 514 shows the amino acid sequence (SEQ ID NO:514) derived from the coding sequence of SEQ ID NO:513 shown in Figure 513.

25 Figure 515 shows a nucleotide sequence (SEQ ID NO:515) of a native sequence PRO1434 cDNA, wherein SEQ ID NO:515 is a clone designated herein as "DNA68818-2536".

Figure 516 shows the amino acid sequence (SEQ ID NO:516) derived from the coding sequence of SEQ ID NO:515 shown in Figure 515.

Figure 517 shows a nucleotide sequence (SEQ ID NO:517) of a native sequence PRO1418 cDNA, wherein SEQ ID NO:517 is a clone designated herein as "DNA68864-1629".

30 Figure 518 shows the amino acid sequence (SEQ ID NO:518) derived from the coding sequence of SEQ ID NO:517 shown in Figure 517.

Figure 519 shows a nucleotide sequence (SEQ ID NO:519) of a native sequence PRO1387 cDNA, wherein SEQ ID NO:519 is a clone designated herein as "DNA68872-1620".

35 Figure 520 shows the amino acid sequence (SEQ ID NO:520) derived from the coding sequence of SEQ ID NO:519 shown in Figure 519.

Figure 521 shows a nucleotide sequence (SEQ ID NO:521) of a native sequence PRO1384 cDNA, wherein SEQ ID NO:521 is a clone designated herein as "DNA71159-1617".

Figure 522 shows the amino acid sequence (SEQ ID NO:522) derived from the coding sequence of SEQ ID NO:521 shown in Figure 521.

Figure 523 shows a nucleotide sequence (SEQ ID NO:523) of a native sequence PRO1565 cDNA, wherein SEQ ID NO:523 is a clone designated herein as "DNA73727-1673".

5 Figure 524 shows the amino acid sequence (SEQ ID NO:524) derived from the coding sequence of SEQ ID NO:523 shown in Figure 523.

Figure 525 shows a nucleotide sequence (SEQ ID NO:525) of a native sequence PRO1474 cDNA, wherein SEQ ID NO:525 is a clone designated herein as "DNA73739-1645".

Figure 526 shows the amino acid sequence (SEQ ID NO:526) derived from the coding sequence of SEQ ID NO:525 shown in Figure 525.

10 Figure 527 shows a nucleotide sequence (SEQ ID NO:527) of a native sequence PRO1917 cDNA, wherein SEQ ID NO:527 is a clone designated herein as "DNA76400-2528".

Figure 528 shows the amino acid sequence (SEQ ID NO:528) derived from the coding sequence of SEQ ID NO:527 shown in Figure 527.

15 Figure 529 shows a nucleotide sequence (SEQ ID NO:529) of a native sequence PRO1787 cDNA, wherein SEQ ID NO:529 is a clone designated herein as "DNA76510-2504".

Figure 530 shows the amino acid sequence (SEQ ID NO:530) derived from the coding sequence of SEQ ID NO:529 shown in Figure 529.

Figure 531 shows a nucleotide sequence (SEQ ID NO:531) of a native sequence PRO1556 cDNA, wherein SEQ ID NO:531 is a clone designated herein as "DNA76529-1666".

20 Figure 532 shows the amino acid sequence (SEQ ID NO:532) derived from the coding sequence of SEQ ID NO:531 shown in Figure 531.

Figure 533 shows a nucleotide sequence (SEQ ID NO:533) of a native sequence PRO1561 cDNA, wherein SEQ ID NO:533 is a clone designated herein as "DNA76538-1670".

25 Figure 534 shows the amino acid sequence (SEQ ID NO:534) derived from the coding sequence of SEQ ID NO:533 shown in Figure 533.

Figure 535 shows a nucleotide sequence (SEQ ID NO:535) of a native sequence PRO1693 cDNA, wherein SEQ ID NO:535 is a clone designated herein as "DNA77301-1708".

Figure 536 shows the amino acid sequence (SEQ ID NO:536) derived from the coding sequence of SEQ ID NO:535 shown in Figure 535.

30 Figure 537 shows a nucleotide sequence (SEQ ID NO:537) of a native sequence PRO1868 cDNA, wherein SEQ ID NO:537 is a clone designated herein as "DNA77624-2515".

Figure 538 shows the amino acid sequence (SEQ ID NO:538) derived from the coding sequence of SEQ ID NO:537 shown in Figure 537.

35 Figure 539 shows a nucleotide sequence (SEQ ID NO:539) of a native sequence PRO1890 cDNA, wherein SEQ ID NO:539 is a clone designated herein as "DNA79230-2525".

Figure 540 shows the amino acid sequence (SEQ ID NO:540) derived from the coding sequence of SEQ ID NO:539 shown in Figure 539.

Figure 541 shows a nucleotide sequence (SEQ ID NO:541) of a native sequence PRO1887 cDNA, wherein SEQ ID NO:541 is a clone designated herein as "DNA79862-2522".

Figure 542 shows the amino acid sequence (SEQ ID NO:542) derived from the coding sequence of SEQ ID NO:541 shown in Figure 541.

Figure 543 shows a nucleotide sequence (SEQ ID NO:543) of a native sequence PRO4353 cDNA, wherein SEQ ID NO:543 is a clone designated herein as "DNA80145-2594".

Figure 544 shows the amino acid sequence (SEQ ID NO:544) derived from the coding sequence of SEQ ID NO:543 shown in Figure 543.

Figure 545 shows a nucleotide sequence (SEQ ID NO:545) of a native sequence PRO1801 cDNA, wherein SEQ ID NO:545 is a clone designated herein as "DNA83500-2506".

Figure 546 shows the amino acid sequence (SEQ ID NO:546) derived from the coding sequence of SEQ ID NO:545 shown in Figure 545.

Figure 547 shows a nucleotide sequence (SEQ ID NO:547) of a native sequence PRO4357 cDNA, wherein SEQ ID NO:547 is a clone designated herein as "DNA84917-2597".

Figure 548 shows the amino acid sequence (SEQ ID NO:548) derived from the coding sequence of SEQ ID NO:547 shown in Figure 547.

Figure 549 shows a nucleotide sequence (SEQ ID NO:549) of a native sequence PRO4302 cDNA, wherein SEQ ID NO:549 is a clone designated herein as "DNA92218-2554".

Figure 550 shows the amino acid sequence (SEQ ID NO:550) derived from the coding sequence of SEQ ID NO:549 shown in Figure 549.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

I. Definitions

The terms "PRO polypeptide" and "PRO" as used herein and when immediately followed by a numerical designation refer to various polypeptides, wherein the complete designation (i.e., PRO/number) refers to specific polypeptide sequences as described herein. The terms "PRO/number polypeptide" and "PRO/number" wherein the term "number" is provided as an actual numerical designation as used herein encompass native sequence polypeptides and polypeptide variants (which are further defined herein). The PRO polypeptides described herein may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods. The term "PRO polypeptide" refers to each individual PRO/number polypeptide disclosed herein. All disclosures in this specification which refer to the "PRO polypeptide" refer to each of the polypeptides individually as well as jointly. For example, descriptions of the preparation of, purification of, derivation of, formation of antibodies to or against, administration of, compositions containing, treatment of a disease with, etc., pertain to each polypeptide of the invention individually. The term "PRO polypeptide" also includes variants of the PRO/number polypeptides disclosed herein.

A "native sequence PRO polypeptide" comprises a polypeptide having the same amino acid sequence as the corresponding PRO polypeptide derived from nature. Such native sequence PRO polypeptides can be

isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence PRO polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of the specific PRO polypeptide (*e.g.*, an extracellular domain sequence), naturally-occurring variant forms (*e.g.*, alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In various embodiments of the invention, the native sequence PRO polypeptides disclosed herein are mature or full-length native sequence polypeptides comprising the full-length amino acids sequences shown in the accompanying figures. Start and stop codons are shown in bold font and underlined in the figures. However, while the PRO polypeptide disclosed in the accompanying figures are shown to begin with methionine residues designated herein as amino acid position 1 in the figures, it is conceivable and possible that other methionine residues located either upstream or downstream from the amino acid position 1 in the figures may be employed as the starting amino acid residue for the PRO polypeptides.

The PRO polypeptide "extracellular domain" or "ECD" refers to a form of the PRO polypeptide which is essentially free of the transmembrane and cytoplasmic domains. Ordinarily, a PRO polypeptide ECD will have less than 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than 0.5% of such domains. It will be understood that any transmembrane domains identified for the PRO polypeptides of the present invention are identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified herein. Optionally, therefore, an extracellular domain of a PRO polypeptide may contain from about 5 or fewer amino acids on either side of the transmembrane domain/extracellular domain boundary as identified in the Examples or specification and such polypeptides, with or without the associated signal peptide, and nucleic acid encoding them, are contemplated by the present invention.

The approximate location of the "signal peptides" of the various PRO polypeptides disclosed herein are shown in the present specification and/or the accompanying figures. It is noted, however, that the C-terminal boundary of a signal peptide may vary, but most likely by no more than about 5 amino acids on either side of the signal peptide C-terminal boundary as initially identified herein, wherein the C-terminal boundary of the signal peptide may be identified pursuant to criteria routinely employed in the art for identifying that type of amino acid sequence element (*e.g.*, Nielsen et al., Prot. Eng. 10:1-6 (1997) and von Heinje et al., Nucl. Acids. Res. 14:4683-4690 (1986)). Moreover, it is also recognized that, in some cases, cleavage of a signal sequence from a secreted polypeptide is not entirely uniform, resulting in more than one secreted species. These mature polypeptides, where the signal peptide is cleaved within no more than about 5 amino acids on either side of the C-terminal boundary of the signal peptide as identified herein, and the polynucleotides encoding them, are contemplated by the present invention.

"PRO polypeptide variant" means an active PRO polypeptide as defined above or below having at least about 80% amino acid sequence identity with a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Such PRO polypeptide variants include, for

instance, PRO polypeptides wherein one or more amino acid residues are added, or deleted, at the N- or C-terminus of the full-length native amino acid sequence. Ordinarily, a PRO polypeptide variant will have at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, PRO variant polypeptides are at least about 10 amino acids in length, alternatively at least about 20 amino acids in length, alternatively at least about 30 amino acids in length, alternatively at least about 40 amino acids in length, alternatively at least about 50 amino acids in length, alternatively at least about 60 amino acids in length, alternatively at least about 70 amino acids in length, alternatively at least about 80 amino acids in length, alternatively at least about 90 amino acids in length, alternatively at least about 100 amino acids in length, alternatively at least about 150 amino acids in length, alternatively at least about 200 amino acids in length, alternatively at least about 300 amino acids in length, or more.

"Percent (%) amino acid sequence identity" with respect to the PRO polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific PRO polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly

available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

In situations where ALIGN-2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

$$100 \text{ times the fraction } X/Y$$

where X is the number of amino acid residues scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A. As examples of % amino acid sequence identity calculations using this method, Tables 2 and 3 demonstrate how to calculate the % amino acid sequence identity of the amino acid sequence designated "Comparison Protein" to the amino acid sequence designated "PRO", wherein "PRO" represents the amino acid sequence of a hypothetical PRO polypeptide of interest, "Comparison Protein" represents the amino acid sequence of a polypeptide against which the "PRO" polypeptide of interest is being compared, and "X," "Y" and "Z" each represent different hypothetical amino acid residues.

Unless specifically stated otherwise, all % amino acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program. However, % amino acid sequence identity values may also be obtained as described below by using the WU-BLAST-2 computer program (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. When WU-BLAST-2 is employed, a % amino acid sequence identity value is determined by dividing (a) the number of matching identical amino acid residues between the amino acid sequence of the PRO polypeptide of interest having a sequence derived from the native PRO polypeptide and the comparison amino acid sequence of interest (i.e., the sequence against which the PRO polypeptide of interest is being compared which may be a PRO variant polypeptide) as determined by WU-BLAST-2 by (b) the total number of amino acid residues of the PRO polypeptide of interest. For example, in the statement "a polypeptide comprising an the amino acid sequence A which has or having at least 80% amino acid sequence identity to the amino acid sequence B", the amino acid sequence A is the comparison amino acid sequence of interest and the amino acid sequence B is the amino acid sequence of the PRO polypeptide of interest.

Percent amino acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997)). The NCBI-BLAST2 sequence

comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov> or otherwise obtained from the National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

$$100 \text{ times the fraction } X/Y$$

where X is the number of amino acid residues scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A.

"PRO variant polynucleotide" or "PRO variant nucleic acid sequence" means a nucleic acid molecule which encodes an active PRO polypeptide as defined below and which has at least about 80% nucleic acid sequence identity with a nucleotide acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, a PRO variant polynucleotide will have at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity with a nucleic acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal sequence, as disclosed herein or any other fragment of a full-

length PRO polypeptide sequence as disclosed herein. Variants do not encompass the native nucleotide sequence.

Ordinarily, PRO variant polynucleotides are at least about 30 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 120 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 180 nucleotides in length, alternatively at least about 210 nucleotides in length, alternatively at least about 240 nucleotides in length, alternatively at least about 270 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 900 nucleotides in length, or more.

"Percent (%) nucleic acid sequence identity" with respect to PRO-encoding nucleic acid sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the PRO nucleic acid sequence of interest, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. For purposes herein, however, % nucleic acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

In situations where ALIGN-2 is employed for nucleic acid sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

$$100 \text{ times the fraction } W/Z$$

where W is the number of nucleotides scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C. As examples of % nucleic acid sequence identity calculations, Tables 4 and 5, demonstrate how to calculate the % nucleic acid sequence identity of the nucleic acid sequence designated "Comparison DNA" to the nucleic acid sequence designated "PRO-DNA", wherein "PRO-DNA" represents a hypothetical PRO-encoding nucleic

acid sequence of interest, "Comparison DNA" represents the nucleotide sequence of a nucleic acid molecule against which the "PRO-DNA" nucleic acid molecule of interest is being compared, and "N", "L" and "V" each represent different hypothetical nucleotides.

Unless specifically stated otherwise, all % nucleic acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program. However, % nucleic acid sequence identity values may also be obtained as described below by using the WU-BLAST-2 computer program (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. When WU-BLAST-2 is employed, a % nucleic acid sequence identity value is determined by dividing (a) the number of matching identical nucleotides between the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest having a sequence derived from the native sequence PRO polypeptide-encoding nucleic acid and the comparison nucleic acid molecule of interest (i.e., the sequence against which the PRO polypeptide-encoding nucleic acid molecule of interest is being compared which may be a variant PRO polynucleotide) as determined by WU-BLAST-2 by (b) the total number of nucleotides of the PRO polypeptide-encoding nucleic acid molecule of interest. For example, in the statement "an isolated nucleic acid molecule comprising a nucleic acid sequence A which has or having at least 80% nucleic acid sequence identity to the nucleic acid sequence B", the nucleic acid sequence A is the comparison nucleic acid molecule of interest and the nucleic acid sequence B is the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest.

Percent nucleic acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov> or otherwise obtained from the National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

$$100 \text{ times the fraction } W/Z$$

where W is the number of nucleotides scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to

C.

In other embodiments, PRO variant polynucleotides are nucleic acid molecules that encode an active PRO polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding a full-length PRO polypeptide as disclosed herein. PRO variant polypeptides may be those that are encoded by a PRO variant polynucleotide.

5 "Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least
10 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the PRO polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

15 An "isolated" PRO polypeptide-encoding nucleic acid or other polypeptide-encoding nucleic acid is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the polypeptide-encoding nucleic acid. An isolated polypeptide-encoding nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated polypeptide-encoding nucleic acid molecules therefore are distinguished from the specific polypeptide-
20 encoding nucleic acid molecule as it exists in natural cells. However, an isolated polypeptide-encoding nucleic acid molecule includes polypeptide-encoding nucleic acid molecules contained in cells that ordinarily express the polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

25 The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

30 Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is
35 accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

The term "antibody" is used in the broadest sense and specifically covers, for example, single anti-PRO monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies), anti-PRO antibody compositions with polyeptopic specificity, single chain anti-PRO antibodies, and fragments of anti-PRO antibodies (see below). The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

"Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

"Stringent conditions" or "high stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

"Moderately stringent conditions" may be identified as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising a PRO polypeptide fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not

substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

As used herein, the term "immunoadhesin" designates antibody-like molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site of an antibody (i.e., is "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD or IgM.

"Active" or "activity" for the purposes herein refers to form(s) of a PRO polypeptide which retain a biological and/or an immunological activity of native or naturally-occurring PRO, wherein "biological" activity refers to a biological function (either inhibitory or stimulatory) caused by a native or naturally-occurring PRO other than the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO and an "immunological" activity refers to the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO.

The term "antagonist" is used in the broadest sense, and includes any molecule that partially or fully blocks, inhibits, or neutralizes a biological activity of a native PRO polypeptide disclosed herein. In a similar manner, the term "agonist" is used in the broadest sense and includes any molecule that mimics a biological activity of a native PRO polypeptide disclosed herein. Suitable agonist or antagonist molecules specifically include agonist or antagonist antibodies or antibody fragments, fragments or amino acid sequence variants of native PRO polypeptides, peptides, antisense oligonucleotides, small organic molecules, etc. Methods for identifying agonists or antagonists of a PRO polypeptide may comprise contacting a PRO polypeptide with a candidate agonist or antagonist molecule and measuring a detectable change in one or more biological activities normally associated with the PRO polypeptide.

"Treatment" refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathologic condition or disorder. Those in need of treatment include those already with the disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented.

"Chronic" administration refers to administration of the agent(s) in a continuous mode as opposed to an acute mode, so as to maintain the initial therapeutic effect (activity) for an extended period of time. "Intermittent" administration is treatment that is not consecutively done without interruption, but rather is cyclic in nature.

"Mammal" for purposes of treatment refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, etc. Preferably, the mammal is human.

Administration "in combination with" one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

"Carriers" as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are nontoxic to the cell or mammal being exposed thereto at the dosages and concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, polyethylene glycol (PEG), and PLURONICS™.

"Antibody fragments" comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')₂, and Fv fragments; diabodies; linear antibodies (Zapata et al., Protein Eng. 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, each with a single antigen-binding site, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. Pepsin treatment yields an F(ab')₂ fragment that has two antigen-combining sites and is still capable of cross-linking antigen.

"Fv" is the minimum antibody fragment which contains a complete antigen-recognition and -binding site. This region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. It is in this configuration that the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the V_H-V_L dimer. Collectively, the six CDRs confer antigen-binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

The Fab fragment also contains the constant domain of the light chain and the first constant domain (CH1) of the heavy chain. Fab fragments differ from Fab' fragments by the addition of a few residues at the carboxy terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. F(ab')₂ antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

The "light chains" of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant domains.

Depending on the amino acid sequence of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of immunoglobulins: IgA, IgD, IgE, IgG, and

IgM, and several of these may be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA, and IgA2.

"Single-chain Fv" or "sFv" antibody fragments comprise the V_H and V_L domains of antibody, wherein these domains are present in a single polypeptide chain. Preferably, the Fv polypeptide further comprises a polypeptide linker between the V_H and V_L domains which enables the sFv to form the desired structure for antigen binding. For a review of sFv, see Pluckthun in The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenberg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994).

The term "diabodies" refers to small antibody fragments with two antigen-binding sites, which fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) in the same polypeptide chain (V_H-V_L). By using a linker that is too short to allow pairing between the two domains on the same chain, the domains are forced to pair with the complementary domains of another chain and create two antigen-binding sites. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and Hollinger et al., Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993).

An "isolated" antibody is one which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or nonreducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody in situ within recombinant cells since at least one component of the antibody's natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

An antibody that "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide is one that binds to that particular polypeptide or epitope on a particular polypeptide without substantially binding to any other polypeptide or polypeptide epitope.

The word "label" when used herein refers to a detectable compound or composition which is conjugated directly or indirectly to the antibody so as to generate a "labeled" antibody. The label may be detectable by itself (e.g. radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition which is detectable.

By "solid phase" is meant a non-aqueous matrix to which the antibody of the present invention can adhere. Examples of solid phases encompassed herein include those formed partially or entirely of glass (e.g., controlled pore glass), polysaccharides (e.g., agarose), polyacrylamides, polystyrene, polyvinyl alcohol and silicones. In certain embodiments, depending on the context, the solid phase can comprise the well of an assay plate; in others it is a purification column (e.g., an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Patent No. 4,275,149.

A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug (such as a PRO polypeptide or antibody thereto) to a mammal. The

components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes.

A "small molecule" is defined herein to have a molecular weight below about 500 Daltons.

An "effective amount" of a polypeptide disclosed herein or an agonist or antagonist thereof is an amount sufficient to carry out a specifically stated purpose. An "effective amount" may be determined empirically and
5 in a routine manner, in relation to the stated purpose.

Table 1

```

/*
 *
 * C-C increased from 12 to 15
 * Z is average of EQ
5  * B is average of ND
 * match with stop is _M; stop-stop = 0; J (joker) match = 0
 */
#define _M      -8      /* value of a match with a stop */

10 int  _day[26][26] = {
/*  A B C D E F G H I J K L M N O P Q R S T U V W X Y Z */
/* A */  { 2, 0, -2, 0, 0, -4, 1, -1, -1, 0, -1, -2, -1, 0, _M, 1, 0, -2, 1, 1, 0, 0, -6, 0, -3, 0},
/* B */  { 0, 3, -4, 3, 2, -5, 0, 1, -2, 0, 0, -3, -2, 2, _M, -1, 1, 0, 0, 0, 0, -2, -5, 0, -3, 1},
/* C */  {-2, -4, 15, -5, -5, -4, -3, -3, -2, 0, -5, -6, -5, -4, _M, -3, -5, -4, 0, -2, 0, -2, -8, 0, 0, -5},
15 /* D */  { 0, 3, -5, 4, 3, -6, 1, 1, -2, 0, 0, -4, -3, 2, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 2},
/* E */  { 0, 2, -5, 3, 4, -5, 0, 1, -2, 0, 0, -3, -2, 1, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 3},
/* F */  {-4, -5, -4, -6, -5, 9, -5, -2, 1, 0, -5, 2, 0, -4, _M, -5, -5, -4, -3, -3, 0, -1, 0, 0, 7, -5},
/* G */  { 1, 0, -3, 1, 0, -5, 5, -2, -3, 0, -2, -4, -3, 0, _M, -1, -1, -3, 1, 0, 0, -1, -7, 0, -5, 0},
/* H */  {-1, 1, -3, 1, 1, -2, -2, 6, -2, 0, 0, -2, -2, 2, _M, 0, 3, 2, -1, -1, 0, -2, -3, 0, 0, 2},
20 /* I */  {-1, -2, -2, -2, -2, 1, -3, -2, 5, 0, -2, 2, 2, -2, _M, -2, -2, -2, -1, 0, 0, 4, -5, 0, -1, -2},
/* J */  { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */  {-1, 0, -5, 0, 0, -5, -2, 0, -2, 0, 5, -3, 0, 1, _M, -1, 1, 3, 0, 0, 0, -2, -3, 0, -4, 0},
/* L */  {-2, -3, -6, -4, -3, 2, -4, -2, 2, 0, -3, 6, 4, -3, _M, -3, -2, -3, -3, -1, 0, 2, -2, 0, -1, -2},
/* M */  {-1, -2, -5, -3, -2, 0, -3, -2, 2, 0, 0, 4, 6, -2, _M, -2, -1, 0, -2, -1, 0, 2, -4, 0, -2, -1},
25 /* N */  { 0, 2, -4, 2, 1, -4, 0, 2, -2, 0, 1, -3, -2, 2, _M, -1, 1, 0, 1, 0, 0, -2, -4, 0, -2, 1},
/* O */  {_M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M},
/* P */  { 1, -1, -3, -1, -1, -5, -1, 0, -2, 0, -1, -3, -2, -1, _M, 6, 0, 0, 1, 0, 0, -1, -6, 0, -5, 0},
/* Q */  { 0, 1, -5, 2, 2, -5, -1, 3, -2, 0, 1, -2, -1, 1, _M, 0, 4, 1, -1, -1, 0, -2, -5, 0, -4, 3},
/* R */  {-2, 0, -4, -1, -1, -4, -3, 2, -2, 0, 3, -3, 0, 0, _M, 0, 1, 6, 0, -1, 0, -2, 2, 0, -4, 0},
30 /* S */  { 1, 0, 0, 0, 0, -3, 1, -1, -1, 0, 0, -3, -2, 1, _M, 1, -1, 0, 2, 1, 0, -1, -2, 0, -3, 0},
/* T */  { 1, 0, -2, 0, 0, -3, 0, -1, 0, 0, 0, -1, -1, 0, _M, 0, -1, -1, 1, 3, 0, 0, -5, 0, -3, 0},
/* U */  { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */  { 0, -2, -2, -2, -2, -1, -1, -2, 4, 0, -2, 2, 2, -2, _M, -1, -2, -2, -1, 0, 0, 4, -6, 0, -2, -2},
/* W */  {-6, -5, -8, -7, -7, 0, -7, -3, -5, 0, -3, -2, -4, -4, _M, -6, -5, 2, -2, -5, 0, -6, 17, 0, 0, -6},
35 /* X */  { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */  {-3, -3, 0, -4, -4, 7, -5, 0, -1, 0, -4, -1, -2, -2, _M, -5, -4, -4, -3, -3, 0, -2, 0, 0, 10, -4},
/* Z */  { 0, 1, -5, 2, 3, -5, 0, 2, -2, 0, 0, -2, -1, 1, _M, 0, 3, 0, 0, 0, 0, -2, -6, 0, -4, 4}
};

40

45

50

55

```


Table 1 (cont')

```

/*
*/
#include <stdio.h>
#include <ctype.h>

5  #define MAXJMP      16      /* max jumps in a diag */
   #define MAXGAP      24      /* don't continue to penalize gaps larger than this */
   #define JMPS        1024    /* max jmps in an path */
10  #define MX          4       /* save if there's at least MX-1 bases since last jmp */

   #define DMAT         3       /* value of matching bases */
   #define DMIS         0       /* penalty for mismatched bases */
   #define DINS0        8       /* penalty for a gap */
   #define DINS1        1       /* penalty per base */
15  #define PINS0        8       /* penalty for a gap */
   #define PINS1        4       /* penalty per residue */

struct jmp {
20     short      n[MAXJMP];    /* size of jmp (neg for del) */
     unsigned short x[MAXJMP]; /* base no. of jmp in seq x */
}; /* limits seq to 2^16 - 1 */

struct diag {
25     int         score;        /* score at last jmp */
     long        offset;        /* offset of prev block */
     short       jmp;          /* current jmp index */
     struct jmp   jp;           /* list of jmps */
};

30 struct path {
     int         spc;           /* number of leading spaces */
     short       n[JMPS];      /* size of jmp (gap) */
     int         x[JMPS];      /* loc of jmp (last elem before gap) */
};

35 char      *ofile;           /* output file name */
char      *name[2];           /* seq names: getseqs() */
char      *prog;              /* prog name for err msgs */
char      *seq[2];            /* seqs: getseqs() */
40 int      dmax;              /* best diag: nw() */
int      dmax0;              /* final diag */
int      dna;                 /* set if dna: main() */
int      endgaps;             /* set if penalizing end gaps */
int      gapx, gapy;          /* total gaps in seqs */
45 int      len0, len1;        /* seq lens */
int      ngapx, ngapy;        /* total size of gaps */
int      smax;                /* max score: nw() */
int      *xbm;                /* bitmap for matching */
long     offset;              /* current offset in jmp file */
50 struct    diag      *dx;     /* holds diagonals */
struct    path      pp[2];     /* holds path for seqs */

char      *calloc(), *malloc(), *index(), *strcpy();
55 char      *getseq(), *g_calloc();

```

60

Table 1 (cont')

```

/* Needleman-Wunsch alignment program
*
* usage: progs file1 file2
* where file1 and file2 are two dna or two protein sequences.
5  * The sequences can be in upper- or lower-case and may contain ambiguity
* Any lines beginning with ';', '>' or '<' are ignored
* Max file length is 65535 (limited by unsigned short x in the jmp struct)
* A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
10  * Output is in the file "align.out"
*
* The program may create a tmp file in /tmp to hold info about traceback.
* Original version developed under BSD 4.3 on a vax 8650
*/
#include "nw.h"
15 #include "day.h"

static  _dbval[26] = {
        1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
20 };

static  _pbval[26] = {
        1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
        128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
25 1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
        1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)                                main
30      int      ac;
      char      *av[];
{
      prog = av[0];
      if (ac != 3) {
35          fprintf(stderr, "usage: %s file1 file2\n", prog);
          fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
          fprintf(stderr, "The sequences can be in upper- or lower-case\n");
          fprintf(stderr, "Any lines beginning with ';', '>' or '<' are ignored\n");
          fprintf(stderr, "Output is in the file \"align.out\"\n");
          exit(1);
40      }
      namex[0] = av[1];
      namex[1] = av[2];
      seqx[0] = getseq(namex[0], &len0);
      seqx[1] = getseq(namex[1], &len1);
45      xbm = (dna)? _dbval : _pbval;

      endgaps = 0;                          /* 1 to penalize endgaps */
      ofile = "align.out";                  /* output file */

50      nw();                                /* fill in the matrix, get the possible jmps */
      readjmps();                           /* get the actual jmps */
      print();                              /* print stats, alignment */

      cleanup(0);                          /* unlink any tmp files */
55  }

```

Table 1 (cont')

```

/* do the alignment, return best score: main()
* dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
* pro: PAM 250 values
* When scores are equal, we prefer mismatches to any gap, prefer
5 * a new gap to extending an ongoing gap, and prefer a gap in seqx
* to a gap in seq y.
*/

```

```

nw()
{

```

nw

```

10 char      *px, *py;          /* seqs and ptrs */
   int      *ndely, *dely;     /* keep track of dely */
   int      ndelx, delx;       /* keep track of delx */
   int      *tmp;              /* for swapping row0, row1 */
15 int      mis;               /* score for each type */
   int      ins0, ins1;        /* insertion penalties */
   register id;                /* diagonal index */
   register ij;                /* jmp index */
   register *col0, *col1;      /* score for curr, last row */
   register xx, yy;            /* index into seqs */

20 dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

   ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
   dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
25 col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
   col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
   ins0 = (dna)? DINS0 : PINS0;
   ins1 = (dna)? DINS1 : PINS1;

30 smax = -10000;
   if (endgaps) {
       for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
           col0[yy] = dely[yy] = col0[yy-1] - ins1;
           ndely[yy] = yy;
35       }
       col0[0] = 0;          /* Waterman Bull Math Biol 84 */
   }
   else
40       for (yy = 1; yy <= len1; yy++)
           dely[yy] = -ins0;

   /* fill in match matrix
   */
45   for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
       /* initialize first entry in col
       */
       if (endgaps) {
           if (xx == 1)
50               col1[0] = delx = -(ins0+ins1);
           else
               col1[0] = delx = col0[0] - ins1;
           ndelx = xx;
       }
       else {
55           col1[0] = 0;
           delx = -ins0;
           ndelx = 0;
       }

```

60

Table 1 (cont')

...nw

```
for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
```

```
    mis = col0[yy-1];
```

```
    if (dna)
```

```
        mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
```

```
    else
```

```
        mis += _day[*px-'A'][*py-'A'];
```

```
    /* update penalty for del in x seq;
```

```
    * favor new del over ongoing del
```

```
    * ignore MAXGAP if weighting endgaps
```

```
    */
```

```
    if (endgaps || ndely[yy] < MAXGAP) {
```

```
        if (col0[yy] - ins0 >= dely[yy]) {
```

```
            dely[yy] = col0[yy] - (ins0+ins1);
```

```
            ndely[yy] = 1;
```

```
        } else {
```

```
            dely[yy] -= ins1;
```

```
            ndely[yy]++;
```

```
        }
```

```
    } else {
```

```
        if (col0[yy] - (ins0+ins1) >= dely[yy]) {
```

```
            dely[yy] = col0[yy] - (ins0+ins1);
```

```
            ndely[yy] = 1;
```

```
        } else
```

```
            ndely[yy]++;
```

```
    }
```

```
    /* update penalty for del in y seq;
```

```
    * favor new del over ongoing del
```

```
    */
```

```
    if (endgaps || ndelx < MAXGAP) {
```

```
        if (col1[yy-1] - ins0 >= delx) {
```

```
            delx = col1[yy-1] - (ins0+ins1);
```

```
            ndelx = 1;
```

```
        } else {
```

```
            delx -= ins1;
```

```
            ndelx++;
```

```
        }
```

```
    } else {
```

```
        if (col1[yy-1] - (ins0+ins1) >= delx) {
```

```
            delx = col1[yy-1] - (ins0+ins1);
```

```
            ndelx = 1;
```

```
        } else
```

```
            ndelx++;
```

```
    }
```

```
    /* pick the maximum score; we're favoring
```

```
    * mis over any del and delx over dely
```

```
    */
```

Table 1 (cont')

...nw

```

id = xx - yy + len1 - 1;
if (mis >= delx && mis >= dely[yy])
    coll[yy] = mis;
5   else if (delx >= dely[yy]) {
        coll[yy] = delx;
        ij = dx[id].ijmp;
        if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
10      && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
            dx[id].ijmp++;
            if (++ij >= MAXJMP) {
                writejumps(id);
                ij = dx[id].ijmp = 0;
                dx[id].offset = offset;
15      offset += sizeof(struct jmp) + sizeof(offset);
            }
        }
        dx[id].jp.n[ij] = ndelx;
        dx[id].jp.x[ij] = xx;
20      dx[id].score = delx;
    }
    else {
        coll[yy] = dely[yy];
        ij = dx[id].ijmp;
25      if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
        && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
            dx[id].ijmp++;
            if (++ij >= MAXJMP) {
                writejumps(id);
                ij = dx[id].ijmp = 0;
                dx[id].offset = offset;
30      offset += sizeof(struct jmp) + sizeof(offset);
            }
        }
        dx[id].jp.n[ij] = -ndely[yy];
        dx[id].jp.x[ij] = xx;
        dx[id].score = dely[yy];
    }
    if (xx == len0 && yy < len1) {
40      /* last col
        */
        if (endgaps)
            coll[yy] -= ins0+ins1*(len1-yy);
        if (coll[yy] > smax) {
45      smax = coll[yy];
            dmax = id;
        }
    }
}
50  if (endgaps && xx < len0)
    coll[yy-1] -= ins0+ins1*(len0-xx);
    if (coll[yy-1] > smax) {
        smax = coll[yy-1];
        dmax = id;
55  }
    tmp = col0; col0 = coll; coll = tmp;
}
(void) free((char *)ndely);
(void) free((char *)dely);
(void) free((char *)col0);
60  (void) free((char *)coll);
    }

```

Table 1 (cont')

```

/*
 *
 * print() -- only routine visible outside this module
 *
5  * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[]: print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
10 * nums() -- put out a number line: dumpblock()
 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() -- put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */

15 #include "nw.h"

#define SPC      3
#define P_LINE   256 /* maximum output line */
#define P_SPC    3    /* space between name or num and seq */

20 extern _day[26][26];
int      olen;        /* set output line length */
FILE     *fx;         /* output file */

25 print()
{
    int      lx, ly, firstgap, lastgap; /* overlap */

    if ((fx = fopen(ofile, "w")) == 0) {
30         fprintf(stderr, "%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
    fprintf(fx, "< first sequence: %s (length = %d)\n", namex[0], len0);
    fprintf(fx, "< second sequence: %s (length = %d)\n", namex[1], len1);
35     olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax < len1 - 1) { /* leading gap in x */
40         pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
45         pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
    if (dmax0 < len0 - 1) { /* trailing gap in x */
50         lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
55     getmat(lx, ly, firstgap, lastgap);
    pr_align();
}

60

```

print

Table 1 (cont')

getmat

```

/*
 * trace back the best path, count matches
 */
static
5 getmat(lx, ly, firstgap, lastgap)
    int    lx, ly;          /* "core" (minus endgaps) */
    int    firstgap, lastgap; /* leading trailing overlap */
{
    int      nm, i0, i1, siz0, siz1;
10    char    outx[32];
    double   pct;
    register n0, n1;
    register char *p0, *p1;

15    /* get total matches, score
    */
    i0 = i1 = siz0 = siz1 = 0;
    p0 = seqx[0] + pp[1].spc;
    p1 = seqx[1] + pp[0].spc;
20    n0 = pp[1].spc + 1;
    n1 = pp[0].spc + 1;

    nm = 0;
    while ( *p0 && *p1 ) {
25        if (siz0) {
            p1++;
            n1++;
            siz0--;
        }
        else if (siz1) {
            p0++;
            n0++;
            siz1--;
        }
35        else {
            if (xbm[*p0-'A']&xbm[*p1-'A'])
                nm++;
            if (n0++ == pp[0].x[i0])
                siz0 = pp[0].n[i0++];
40            if (n1++ == pp[1].x[i1])
                siz1 = pp[1].n[i1++];
            p0++;
            p1++;
        }
45    }

    /* pct homology:
    * if penalizing endgaps, base is the shorter seq
    * else, knock off overhangs and take shorter core
    */
50    if (endgaps)
        lx = (len0 < len1)? len0 : len1;
    else
        lx = (lx < ly)? lx : ly;
55    pct = 100.*((double)nm)/(double)lx;
    fprintf(fx, "\n");
    fprintf(fx, "< %d match%s in an overlap of %d: %.2f percent similarity\n",
        nm, (nm == 1)? "" : "es", lx, pct);
60

```

Table 1 (cont')

```

fprintf(fx, "< gaps in first sequence: %d", gapx);
if (gapx) {
    (void) sprintf(outx, " (%d %s%s)",
        ngapx, (dna)? "base": "residue", (ngapx == 1)? "": "s");
    fprintf(fx, "%s", outx);

    fprintf(fx, ", gaps in second sequence: %d", gapy);
    if (gapy) {
        (void) sprintf(outx, " (%d %s%s)",
            ngapy, (dna)? "base": "residue", (ngapy == 1)? "": "s");
        fprintf(fx, "%s", outx);
    }
    if (dna)
        fprintf(fx,
            "\n<score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
            smax, DMAT, DMIS, DINS0, DINS1);
    else
        fprintf(fx,
            "\n<score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
            smax, PINS0, PINS1);
    if (endgaps)
        fprintf(fx,
            "<endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
            firstgap, (dna)? "base": "residue", (firstgap == 1)? "": "s",
            lastgap, (dna)? "base": "residue", (lastgap == 1)? "": "s");
    else
        fprintf(fx, "<endgaps not penalized\n");
}

static      nm;          /* matches in core -- for checking */
static      lmax;        /* lengths of stripped file names */
static      ij[2];       /* jmp index for a path */
static      nc[2];       /* number at start of current line */
static      ni[2];       /* current elem number -- for gapping */
static      siz[2];
static char *ps[2];      /* ptr to current element */
static char *po[2];      /* ptr to next output char slot */
static char out[2][P_LINE]; /* output line */
static char star[P_LINE]; /* set by stars() */

/*
 * print alignment of described in struct path pp[]
 */
static
pr_align()
{
    int      nn;          /* char count */
    int      more;
    register i;

    for (i = 0, lmax = 0; i < 2; i++) {
        nn = stripname(name[i]);
        if (nn > lmax)
            lmax = nn;

        nc[i] = 1;
        ni[i] = 1;
        siz[i] = ij[i] = 0;
        ps[i] = seqx[i];
        po[i] = out[i];
    }
}

```

...getmat

pr_align

Table 1 (cont')

...pr_align

```

for (nn = nm = 0, more = 1; more; ) {
    for (i = more = 0; i < 2; i++) {
        /*
5         * do we have more of this sequence?
        */
        if (!*ps[i])
            continue;

10        more++;

        if (pp[i].spc) { /* leading space */
            *po[i]++ = ' ';
            pp[i].spc--;
15        }
        else if (siz[i]) { /* in a gap */
            *po[i]++ = '-';
            siz[i]--;
20        }
        else { /* we're putting a seq element
            */
            *po[i] = *ps[i];
            if (islower(*ps[i]))
                *ps[i] = toupper(*ps[i]);
25            po[i]++;
            ps[i]++;

            /*
            * are we at next gap for this seq?
            */
            if (ni[i] == pp[i].x[ij[i]]) {
                /*
                * we need to merge all gaps
                * at this location
                */
35                siz[i] = pp[i].n[ij[i]++];
                while (ni[i] == pp[i].x[ij[i]])
                    siz[i] += pp[i].n[ij[i]++];
            }
            ni[i]++;
40        }
    }
    if (++nn == olen || !more && nn) {
        dumpblock();
        for (i = 0; i < 2; i++)
            po[i] = out[i];
        nn = 0;
    }
50 }

/*
 * dump a block of lines, including numbers, stars: pr_align()
 */
55 static
dumpblock()
{
    register i;

60    for (i = 0; i < 2; i++)
        *po[i]-- = '\0';

```

dumpblock

Table 1 (cont')

...dumpblock

```

5      (void) putc('\n', fx);
      for (i = 0; i < 2; i++) {
          if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
              if (i == 0)
                  nums(i);
              if (i == 0 && *out[1])
                  stars();
10         putline(i);
              if (i == 0 && *out[1])
                  fprintf(fx, star);
              if (i == 1)
                  nums(i);
15     }
    }
}

/*
20  * put out a number line: dumpblock()
   */
static
nums(ix)
25 {
    int      ix;      /* index in out[] holding seq line */
    char      nline[P_LINE];
    register  i, j;
    register char *pn, *px, *py;

30     for (pn = nline, i = 0; i < lmax + P_SPC; i++, pn++)
        *pn = ' ';
    for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
        if (*py == ' ' || *py == '-')
            *pn = ' ';
35         else {
            if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                j = (i < 0)? -i : i;
                for (px = pn; j /= 10, px--)
                    *px = j%10 + '0';
40                 if (i < 0)
                    *px = '-';
            }
            else
                *pn = ' ';
45             i++;
        }
    }
    *pn = '\0';
    nc[ix] = i;
50     for (pn = nline; *pn; pn++)
        (void) putc(*pn, fx);
    (void) putc('\n', fx);
}

55  /*
   * put out a line (name, [num], seq, [num]): dumpblock()
   */
static
putline(ix)
60 int      ix;      {

```

nums

putline

Table 1 (cont')

...putline

```

5      int          i;
      register char *px;

      for (px = namex[ix], i = 0; *px && *px != '.'; px++, i++)
          (void) putc(*px, fx);
      for (; i < lmax+P_SPC; i++)
          (void) putc(' ', fx);

10     /* these count from 1:
       * ni[] is current element (from 1)
       * nc[] is number at start of current line
       */
15     for (px = out[ix]; *px; px++)
          (void) putc(*px&0x7F, fx);
      (void) putc('\n', fx);
  }

20  /*
   * put a line of stars (seqs always in out[0], out[1]): dumpblock()
   */
   static
25  stars()
   {
       int          i;
       register char *p0, *p1, cx, *px;

30     if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
        !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
          return;
       px = star;
       for (i = lmax+P_SPC; i; i--)
35         *px++ = ' ';

       for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
           if (isalpha(*p0) && isalpha(*p1)) {
40               if (xbm[*p0-'A']&xbm[*p1-'A']) {
                   cx = '*';
                   nm++;
               }
               else if (!dna && _day[*p0-'A'][*p1-'A'] > 0)
45                   cx = '.';
               else
                   cx = ' ';
           }
           else
50               cx = ' ';
           *px++ = cx;
       }
       *px++ = '\n';
       *px = '\0';
55  }

```

stars

Table 1 (cont')

stripname

```
/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
5 stripname(pn)
    char    *pn;    /* file name (may be path) */
{
    register char    *px, *py;
10     py = 0;
    for (px = pn; *px; px++)
        if (*px == '/')
            py = px + 1;
15     if (py)
        (void) strcpy(pn, py);
    return(strlen(pn));
}
```

Table 1 (cont')

```

/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_calloc() -- calloc() with error checkin
5  * readjumps() -- get the good jumps, from tmp file if necessary
 * writejumps() -- write a filled array of jumps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>

10 char *jname = "/tmp/homgXXXXXX"; /* tmp file for jumps */
FILE *fj;

int cleanup(); /* cleanup tmp file */
15 long lseek();

/*
 * remove any tmp file if we blow
 */
20 cleanup(i)
    int i;
{
    if (fj)
        (void) unlink(jname);
25    exit(i);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
30 char *
getseq(file, len)
35 char *file; /* file name */
    int *len; /* seq len */
{
    char line[1024], *pseq;
    register char *px, *py;
    int natgc, tlen;
    FILE *fp;

    if ((fp = fopen(file, "r")) == 0) {
        fprintf(stderr, "%s: can't read %s\n", prog, file);
        exit(1);
    }
    tlen = natgc = 0;
    while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++)
            if (isupper(*px) || islower(*px))
                tlen++;
    }
    if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr, "%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
    }
    pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
60

```

cleanup

getseq

Table 1 (cont')

...getseq

```

py = pseq + 4;
*len = tlen;
rewind(fp);

5   while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++) {
10          if (isupper(*px))
                *py++ = *px;
            else if (islower(*px))
                *py++ = toupper(*px);
            if (index("ATGCU",*(py-1)))
15              natgc++;
        }
    }
    *py++ = '\0';
    *py = '\0';
20   (void) fclose(fp);
    dna = natgc > (tlen/3);
    return(pseq+4);
}

25 char *
g_alloc(msg, nx, sz)
    char *msg;          /* program, calling routine */
    int nx, sz;          /* number and size of elements */
{
30     char *px, *calloc();

    if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
        if (*msg) {
35             fprintf(stderr, "%s: g_alloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
            exit(1);
        }
    }
    return(px);
}

40 /*
 * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
 */
readjmps()
45 {
    int fd = -1;
    int siz, i0, i1;
    register i, j, xx;

50     if (fj) {
        (void) fclose(fj);
        if ((fd = open(jname, O_RDONLY, 0)) < 0) {
            fprintf(stderr, "%s: can't open() %s\n", prog, jname);
            cleanup(1);
55         }
    }
    for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {
        while (1) {
60             for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                ;

```

g_alloc

readjmps

Table 1 (cont')

...readjumps

```

5      if (j < 0 && dx[dmax].offset && fj) {
        (void) lseek(fd, dx[dmax].offset, 0);
        (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
        (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
        dx[dmax].ijmp = MAXJMP-1;
      }
      else
10         break;
    }
    if (i >= JMPS) {
        fprintf(stderr, "%s: too many gaps in alignment\n", prog);
        cleanup(1);
    }
15    if (j >= 0) {
        siz = dx[dmax].jp.n[j];
        xx = dx[dmax].jp.x[j];
        dmax += siz;
        if (siz < 0) { /* gap in second seq */
20            pp[1].n[i1] = -siz;
            xx += siz;
            /* id = xx - yy + len1 - 1 */
            /*
            pp[1].x[i1] = xx - dmax + len1 - 1;
25            gapy++;
            ngapy -= siz;
            /* ignore MAXGAP when doing endgaps */
            siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
            i1++;
30        }
        else if (siz > 0) { /* gap in first seq */
            pp[0].n[i0] = siz;
            pp[0].x[i0] = xx;
            gapx++;
            ngapx += siz;
35            /* ignore MAXGAP when doing endgaps */
            siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
            i0++;
        }
40    }
    else
        break;
}

45    /* reverse the order of jumps
    */
    for (j = 0, i0--; j < i0; j++, i0--) {
        i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
        i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
50    }
    for (j = 0, i1--; j < i1; j++, i1--) {
        i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
        i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
55    }
    if (fd >= 0)
        (void) close(fd);
    if (fj) {
        (void) unlink(jname);
        fj = 0;
        offset = 0;
60    }
}

```

Table 1 (cont')**writejumps**

```

/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
5  writejumps(ix)
    int    ix;
    {
        char    *mktemp();
10         if (!fj) {
            if (mktemp(jname) < 0) {
                fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
                cleanup(1);
            }
15         if ((fj = fopen(jname, "w")) == 0) {
            fprintf(stderr, "%s: can't write %s\n", prog, jname);
            exit(1);
        }
20         (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
        (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
    }
}

```


Table 2

PRO	XXXXXXXXXXXXXXXXXX	(Length = 15 amino acids)
Comparison Protein	XXXXXXXXYYYYYYY	(Length = 12 amino acids)

5 % amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =

10 5 divided by 15 = 33.3%

Table 3

15 PRO	XXXXXXXXXXX	(Length = 10 amino acids)
Comparison Protein	XXXXXXXXYYYYYYZZYZ	(Length = 15 amino acids)

% amino acid sequence identity =

20 (the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =

5 divided by 10 = 50%

Table 4

PRO-DNA	NNNNNNNNNNNNNN	(Length = 14 nucleotides)
Comparison DNA	NNNNNNLLLLLLLLLL	(Length = 16 nucleotides)

5 % nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =

10 6 divided by 14 = 42.9%

Table 5

15 PRO-DNA	NNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNLLLV	(Length = 9 nucleotides)

% nucleic acid sequence identity =

20 (the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =

4 divided by 12 = 33.3%

II. Compositions and Methods of the Invention

A. Full-Length PRO Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO polypeptides. In particular, cDNAs encoding various PRO polypeptides have been identified and isolated, as disclosed in further detail in the Examples below. It is noted that proteins produced in separate expression rounds may be given different PRO numbers but the UNQ number is unique for any given DNA and the encoded protein, and will not be changed. However, for sake of simplicity, in the present specification the protein encoded by the full length native nucleic acid molecules disclosed herein as well as all further native homologues and variants included in the foregoing definition of PRO, will be referred to as "PRO/number", regardless of their origin or mode of preparation.

As disclosed in the Examples below, various cDNA clones have been deposited with the ATCC. The actual nucleotide sequences of those clones can readily be determined by the skilled artisan by sequencing of the deposited clone using routine methods in the art. The predicted amino acid sequence can be determined from the nucleotide sequence using routine skill. For the PRO polypeptides and encoding nucleic acids described herein, Applicants have identified what is believed to be the reading frame best identifiable with the sequence information available at the time.

B. PRO Polypeptide Variants

In addition to the full-length native sequence PRO polypeptides described herein, it is contemplated that PRO variants can be prepared. PRO variants can be prepared by introducing appropriate nucleotide changes into the PRO DNA, and/or by synthesis of the desired PRO polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

Variations in the native full-length sequence PRO or in various domains of the PRO described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO that results in a change in the amino acid sequence of the PRO as compared with the native sequence PRO. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the PRO with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

PRO polypeptide fragments are provided herein. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full length native protein. Certain fragments lack amino acid residues that are not essential for a desired biological activity of the PRO polypeptide.

PRO fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized. An alternative approach involves generating PRO fragments by enzymatic digestion, e.g., by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the desired fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired polypeptide fragment, by polymerase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, PRO polypeptide fragments share at least one biological and/or immunological activity with the native PRO polypeptide disclosed herein.

In particular embodiments, conservative substitutions of interest are shown in Table 6 under the heading of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions in Table 6, or as further described below in reference to amino acid classes, are introduced and the products screened.

Table 6

	<u>Original Residue</u>	<u>Exemplary Substitutions</u>	<u>Preferred Substitutions</u>
	Ala (A)	val; leu; ile	val
	Arg (R)	lys; gln; asn	lys
25	Asn (N)	gln; his; lys; arg	gln
	Asp (D)	glu	glu
	Cys (C)	ser	ser
	Gln (Q)	asn	asn
	Glu (E)	asp	asp
30	Gly (G)	pro; ala	ala
	His (H)	asn; gln; lys; arg	arg
	Ile (I)	leu; val; met; ala; phe; norleucine	leu
	Leu (L)	norleucine; ile; val; met; ala; phe	ile
35	Lys (K)	arg; gln; asn	arg
	Met (M)	leu; phe; ile	leu
	Phe (F)	leu; val; ile; ala; tyr	leu
	Pro (P)	ala	ala
40	Ser (S)	thr	thr
	Thr (T)	ser	ser
	Trp (W)	tyr; phe	tyr
	Tyr (Y)	trp; phe; thr; ser	phe
	Val (V)	ile; leu; met; phe; ala; norleucine	leu
45			

Substantial modifications in function or immunological identity of the PRO polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

- 5 (1) hydrophobic: norleucine, met, ala, val, leu, ile;
- (2) neutral hydrophilic: cys, ser, thr;
- (3) acidic: asp, glu;
- (4) basic: asn, gln, his, lys, arg;
- (5) residues that influence chain orientation: gly, pro; and
- 10 (6) aromatic: trp, tyr, phe.

Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells et al., Gene, 34:315 (1985)], restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO variant DNA.

20 Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, Science, 244: 1081-1085 (1989)]. Alanine is also

25 typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

C. Modifications of PRO

30 Covalent modifications of PRO are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of the PRO. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO antibodies, and vice-versa. Commonly used

35 crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-

octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimide.

Other modifications include deamidation of glutamyl and asparagyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the native sequence PRO. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

Addition of glycosylation sites to the PRO polypeptide may be accomplished by altering the amino acid sequence. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO (for O-linked glycosylation sites). The PRO amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the PRO polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Removal of carbohydrate moieties present on the PRO polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of PRO comprises linking the PRO polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

The PRO of the present invention may also be modified in a way to form a chimeric molecule comprising PRO fused to another, heterologous polypeptide or amino acid sequence.

In one embodiment, such a chimeric molecule comprises a fusion of the PRO with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the PRO. The presence of such epitope-tagged forms of the PRO can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO to

be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an α -tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

In an alternative embodiment, the chimeric molecule may comprise a fusion of the PRO with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a PRO polypeptide in place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH2 and CH3, or the hinge, CH1, CH2 and CH3 regions of an IgG1 molecule. For the production of immunoglobulin fusions see also US Patent No. 5,428,130 issued June 27, 1995.

D. Preparation of PRO

The description below relates primarily to production of PRO by culturing cells transformed or transfected with a vector containing PRO nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO. For instance, the PRO sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of the PRO may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length PRO.

1. Isolation of DNA Encoding PRO

DNA encoding PRO may be obtained from a cDNA library prepared from tissue believed to possess the PRO mRNA and to express it at a detectable level. Accordingly, human PRO DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO-encoding gene may also be obtained from a genomic library or by known synthetic procedures (e.g., automated nucleic acid synthesis).

Libraries can be screened with probes (such as antibodies to the PRO or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO is to use PCR methodology [Sambrook et al., supra; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ³²P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., supra.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined using methods known in the art and as described herein.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for PRO production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

Methods of eukaryotic cell transfection and prokaryotic cell transformation are known to the ordinarily skilled artisan, for example, CaCl₂, CaPO₄, liposome-mediated and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is generally used for prokaryotes. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transfections have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the

method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

5 Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Other suitable prokaryotic host cells include
10 Enterobacteriaceae such as *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescens*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (e.g., *B. licheniformis* 41P disclosed in DD 266,710 published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant
15 DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype *tonA*; *E. coli* W3110 strain 9E4, which has the complete genotype *tonA ptr3*; *E. coli* W3110 strain 27C7 (ATCC 55,244), which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT kan'*; *E. coli*
20 W3110 strain 37D6, which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan'*; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, e.g., PCR or other nucleic acid polymerase reactions, are suitable.

25 In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for PRO-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, Nature, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Fleer et al., Bio/Technology, 9:968-975 (1991)) such as, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al., J. Bacteriol., 154(2):737-742 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickerhamii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilum* (ATCC 36,906; Van den Berg et al., Bio/Technology, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus*; *yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna et al., J. Basic Microbiol., 28:265-278 [1988]); *Candida*; *Trichoderma reesia* (EP 244,234); *Neurospora crassa* (Case et al., Proc. Natl. Acad. Sci. USA, 76:5259-5263 [1979]); *Schwanniomyces*
30 such as *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, e.g., *Neurospora*, *Penicillium*, *Tolytocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance et al., Biochem. Biophys. Res. Commun., 112:284-289 [1983]; Tilburn et al.,

Gene, 26:205-221 [1983]; Yelton et al., Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, EMBO J., 4:475-479 [1985]). Methylophilic yeasts are suitable herein and include, but are not limited to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis*, and *Rhodotorula*. A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, The Biochemistry of Methylophilic Yeasts, 269 (1982).

5 Suitable host cells for the expression of glycosylated PRO are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J.
10 Gen. Virol., 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

15 3. Selection and Use of a Replicable Vector

The nucleic acid (e.g., cDNA or genomic DNA) encoding PRO may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an
20 appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

25 The PRO may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g.,
30 the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α -factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the
35 protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2 μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

5 Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*.

10 An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7 [Stinchcomb et al., Nature, 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)]. The *trp1* gene
15 provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

Expression and cloning vectors usually contain a promoter operably linked to the PRO-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β -lactamase and lactose promoter systems
20 [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (*trp*) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid promoters such as the *tac* promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding PRO.

25 Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose
30 isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and
35 promoters for use in yeast expression are further described in EP 73,657.

PRO transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July

1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a DNA encoding the PRO by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO in recombinant vertebrate cell culture are described in Gething et al., Nature, 293:620-625 (1981); Mantei et al., Nature, 281:40-46 (1979); EP 117,060; and EP 117,058.

4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO DNA and encoding a specific antibody epitope.

5. Purification of Polypeptide

Forms of PRO may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic cleavage. Cells employed in expression of PRO can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify PRO from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO produced.

E. Uses for PRO

Nucleotide sequences (or their complement) encoding PRO have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. PRO nucleic acid will also be useful for the preparation of PRO polypeptides by the recombinant techniques described herein.

The full-length native sequence PRO gene, or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length PRO cDNA or to isolate still other cDNAs (for instance, those encoding naturally-occurring variants of PRO or PRO from other species) which have a desired sequence identity to the native PRO sequence disclosed herein. Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from at least partially novel regions of the full length native nucleotide sequence wherein those regions may be determined without undue experimentation or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO. By way of example, a screening method will comprise isolating the coding region of the PRO gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ^{32}P or ^{35}S , or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the PRO gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

Any EST sequences disclosed in the present application may similarly be employed as probes, using the methods disclosed herein.

Other useful fragments of the PRO nucleic acids include antisense or sense oligonucleotides comprising a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target PRO mRNA (sense) or PRO DNA (antisense) sequences. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment of the coding region of PRO DNA. Such a fragment generally comprises at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen (Cancer Res. 48:2659, 1988) and van der Krol et al. (BioTechniques 6:958, 1988).

Binding of antisense or sense oligonucleotides to target nucleic acid sequences results in the formation of duplexes that block transcription or translation of the target sequence by one of several means, including enhanced degradation of the duplexes, premature termination of transcription or translation, or by other means.

The antisense oligonucleotides thus may be used to block expression of PRO proteins. Antisense or sense oligonucleotides further comprise oligonucleotides having modified sugar-phosphodiester backbones (or other sugar linkages, such as those described in WO 91/06629) and wherein such sugar linkages are resistant to endogenous nucleases. Such oligonucleotides with resistant sugar linkages are stable *in vivo* (i.e., capable of resisting enzymatic degradation) but retain sequence specificity to be able to bind to target nucleotide sequences.

Other examples of sense or antisense oligonucleotides include those oligonucleotides which are covalently linked to organic moieties, such as those described in WO 90/10048, and other moieties that increases affinity of the oligonucleotide for a target nucleic acid sequence, such as poly-(L-lysine). Further still, intercalating agents, such as ellipticine, and alkylating agents or metal complexes may be attached to sense or antisense oligonucleotides to modify binding specificities of the antisense or sense oligonucleotide for the target nucleotide sequence.

Antisense or sense oligonucleotides may be introduced into a cell containing the target nucleic acid sequence by any gene transfer method, including, for example, CaPO₄-mediated DNA transfection, electroporation, or by using gene transfer vectors such as Epstein-Barr virus. In a preferred procedure, an antisense or sense oligonucleotide is inserted into a suitable retroviral vector. A cell containing the target nucleic acid sequence is contacted with the recombinant retroviral vector, either *in vivo* or *ex vivo*. Suitable retroviral vectors include, but are not limited to, those derived from the murine retrovirus M-MuLV, N2 (a retrovirus derived from M-MuLV), or the double copy vectors designated DCT5A, DCT5B and DCT5C (see WO 90/13641).

Sense or antisense oligonucleotides also may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell.

Alternatively, a sense or an antisense oligonucleotide may be introduced into a cell containing the target nucleic acid sequence by formation of an oligonucleotide-lipid complex, as described in WO 90/10448. The sense or antisense oligonucleotide-lipid complex is preferably dissociated within the cell by an endogenous lipase.

Antisense or sense RNA or DNA molecules are generally at least about 5 bases in length, about 10 bases in length, about 15 bases in length, about 20 bases in length, about 25 bases in length, about 30 bases in length, about 35 bases in length, about 40 bases in length, about 45 bases in length, about 50 bases in length, about 55 bases in length, about 60 bases in length, about 65 bases in length, about 70 bases in length, about 75 bases in length, about 80 bases in length, about 85 bases in length, about 90 bases in length, about 95 bases in length, about 100 bases in length, or more.

The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related PRO coding sequences.

Nucleotide sequences encoding a PRO can also be used to construct hybridization probes for mapping the gene which encodes that PRO and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

When the coding sequences for PRO encode a protein which binds to another protein (example, where the PRO is a receptor), the PRO can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor PRO can be used to isolate correlative ligand(s). Screening assays can be designed to find lead compounds that mimic the biological activity of a native PRO or a receptor for PRO. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

Nucleic acids which encode PRO or its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding PRO. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding PRO introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding PRO. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of

the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of PRO can be used to construct a PRO "knock out" animal which has a defective or altered gene encoding PRO as a result of homologous recombination between the endogenous gene encoding PRO and altered genomic DNA encoding PRO introduced into an embryonic stem cell of the animal. For example, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques. A portion of the genomic DNA encoding PRO can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, *Cell*, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., *Cell*, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the PRO polypeptide.

Nucleic acid encoding the PRO polypeptides may also be used in gene therapy. In gene therapy applications, genes are introduced into cells in order to achieve *in vivo* synthesis of a therapeutically effective genetic product, for example for replacement of a defective gene. "Gene therapy" includes both conventional gene therapy where a lasting effect is achieved by a single treatment, and the administration of gene therapeutic agents, which involves the one time or repeated administration of a therapeutically effective DNA or mRNA. Antisense RNAs and DNAs can be used as therapeutic agents for blocking the expression of certain genes *in vivo*. It has already been shown that short antisense oligonucleotides can be imported into cells where they act as inhibitors, despite their low intracellular concentrations caused by their restricted uptake by the cell membrane. (Zamecnik *et al.*, *Proc. Natl. Acad. Sci. USA* 83:4143-4146 [1986]). The oligonucleotides can be modified to enhance their uptake, e.g. by substituting their negatively charged phosphodiester groups by uncharged groups.

There are a variety of techniques available for introducing nucleic acids into viable cells. The techniques vary depending upon whether the nucleic acid is transferred into cultured cells *in vitro*, or *in vivo* in the cells of the intended host. Techniques suitable for the transfer of nucleic acid into mammalian cells *in vitro* include the use of liposomes, electroporation, microinjection, cell fusion, DEAE-dextran, the calcium phosphate precipitation method, etc. The currently preferred *in vivo* gene transfer techniques include transfection with viral

(typically retroviral) vectors and viral coat protein-liposome mediated transfection (Dzau et al., Trends in Biotechnology 11, 205-210 [1993]). In some situations it is desirable to provide the nucleic acid source with an agent that targets the target cells, such as an antibody specific for a cell surface membrane protein or the target cell, a ligand for a receptor on the target cell, etc. Where liposomes are employed, proteins which bind to a cell surface membrane protein associated with endocytosis may be used for targeting and/or to facilitate uptake, e.g. capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half-life. The technique of receptor-mediated endocytosis is described, for example, by Wu et al., J. Biol. Chem. 262, 4429-4432 (1987); and Wagner et al., Proc. Natl. Acad. Sci. USA 87, 3410-3414 (1990). For review of gene marking and gene therapy protocols see Anderson et al., Science 256, 808-813 (1992).

The PRO polypeptides described herein may also be employed as molecular weight markers for protein electrophoresis purposes and the isolated nucleic acid sequences may be used for recombinantly expressing those markers.

The nucleic acid molecules encoding the PRO polypeptides or fragments thereof described herein are useful for chromosome identification. In this regard, there exists an ongoing need to identify new chromosome markers, since relatively few chromosome marking reagents, based upon actual sequence data are presently available. Each PRO nucleic acid molecule of the present invention can be used as a chromosome marker.

The PRO polypeptides and nucleic acid molecules of the present invention may also be used diagnostically for tissue typing, wherein the PRO polypeptides of the present invention may be differentially expressed in one tissue as compared to another, preferably in a diseased tissue as compared to a normal tissue of the same tissue type. PRO nucleic acid molecules will find use for generating probes for PCR, Northern analysis, Southern analysis and Western analysis.

The PRO polypeptides described herein may also be employed as therapeutic agents. The PRO polypeptides of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby the PRO product hereof is combined in admixture with a pharmaceutically acceptable carrier vehicle. Therapeutic formulations are prepared for storage by mixing the active ingredient having the desired degree of purity with optional physiologically acceptable carriers, excipients or stabilizers (Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)), in the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients or stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone, amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, PLURONIC™ or PEG.

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes, prior to or following lyophilization and reconstitution.

Therapeutic compositions herein generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

The route of administration is in accord with known methods, e.g. injection or infusion by intravenous, intraperitoneal, intracerebral, intramuscular, intraocular, intraarterial or intralesional routes, topical administration, or by sustained release systems.

5 Dosages and desired drug concentrations of pharmaceutical compositions of the present invention may vary depending on the particular use envisioned. The determination of the appropriate dosage or route of administration is well within the skill of an ordinary physician. Animal experiments provide reliable guidance for the determination of effective doses for human therapy. Interspecies scaling of effective doses can be performed following the principles laid down by Mordenti, J. and Chappell, W. "The use of interspecies scaling
10 in toxicokinetics" In Toxicokinetics and New Drug Development, Yacobi et al., Eds., Pergamon Press, New York 1989, pp. 42-96.

When *in vivo* administration of a PRO polypeptide or agonist or antagonist thereof is employed, normal dosage amounts may vary from about 10 ng/kg to up to 100 mg/kg of mammal body weight or more per day, preferably about 1 µg/kg/day to 10 mg/kg/day, depending upon the route of administration. Guidance as to
15 particular dosages and methods of delivery is provided in the literature; see, for example, U.S. Pat. Nos. 4,657,760; 5,206,344; or 5,225,212. It is anticipated that different formulations will be effective for different treatment compounds and different disorders, that administration targeting one organ or tissue, for example, may necessitate delivery in a manner different from that to another organ or tissue.

Where sustained-release administration of a PRO polypeptide is desired in a formulation with release
20 characteristics suitable for the treatment of any disease or disorder requiring administration of the PRO polypeptide, microencapsulation of the PRO polypeptide is contemplated. Microencapsulation of recombinant proteins for sustained release has been successfully performed with human growth hormone (rhGH), interferon-
(rhIFN-), interleukin-2, and MN rgp120. Johnson et al., Nat. Med., 2:795-799 (1996); Yasuda, Biomed. Ther., 27:1221-1223 (1993); Hora et al., Bio/Technology, 8:755-758 (1990); Cleland, "Design and Production
25 of Single Immunization Vaccines Using Polylactide Polyglycolide Microsphere Systems," in Vaccine Design: The Subunit and Adjuvant Approach, Powell and Newman, eds, (Plenum Press: New York, 1995), pp. 439-462; WO 97/03692, WO 96/40072, WO 96/07399; and U.S. Pat. No. 5,654,010.

The sustained-release formulations of these proteins were developed using poly-lactic-coglycolic acid (PLGA) polymer due to its biocompatibility and wide range of biodegradable properties. The degradation
30 products of PLGA, lactic and glycolic acids, can be cleared quickly within the human body. Moreover, the degradability of this polymer can be adjusted from months to years depending on its molecular weight and composition. Lewis, "Controlled release of bioactive agents from lactide/glycolide polymer," in: M. Chasin and R. Langer (Eds.), Biodegradable Polymers as Drug Delivery Systems (Marcel Dekker: New York, 1990), pp. 1-41.

35 This invention encompasses methods of screening compounds to identify those that mimic the PRO polypeptide (agonists) or prevent the effect of the PRO polypeptide (antagonists). Screening assays for antagonist drug candidates are designed to identify compounds that bind or complex with the PRO polypeptides

encoded by the genes identified herein, or otherwise interfere with the interaction of the encoded polypeptides with other cellular proteins. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates.

The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays, and cell-based assays, which are well characterized in the art.

5 All assays for antagonists are common in that they call for contacting the drug candidate with a PRO polypeptide encoded by a nucleic acid identified herein under conditions and for a time sufficient to allow these two components to interact.

10 In binding assays, the interaction is binding and the complex formed can be isolated or detected in the reaction mixture. In a particular embodiment, the PRO polypeptide encoded by the gene identified herein or the drug candidate is immobilized on a solid phase, e.g., on a microtiter plate, by covalent or non-covalent attachments. Non-covalent attachment generally is accomplished by coating the solid surface with a solution of the PRO polypeptide and drying. Alternatively, an immobilized antibody, e.g., a monoclonal antibody, specific for the PRO polypeptide to be immobilized can be used to anchor it to a solid surface. The assay is performed by adding the non-immobilized component, which may be labeled by a detectable label, to the immobilized component, e.g., the coated surface containing the anchored component. When the reaction is complete, the non-reacted components are removed, e.g., by washing, and complexes anchored on the solid surface are detected. When the originally non-immobilized component carries a detectable label, the detection of label immobilized on the surface indicates that complexing occurred. Where the originally non-immobilized component does not carry a label, complexing can be detected, for example, by using a labeled antibody specifically binding the immobilized complex.

20 If the candidate compound interacts with but does not bind to a particular PRO polypeptide encoded by a gene identified herein, its interaction with that polypeptide can be assayed by methods well known for detecting protein-protein interactions. Such assays include traditional approaches, such as, e.g., cross-linking, co-immunoprecipitation, and co-purification through gradients or chromatographic columns. In addition, protein-protein interactions can be monitored by using a yeast-based genetic system described by Fields and co-workers (Fields and Song, Nature (London), 340:245-246 (1989); Chien et al., Proc. Natl. Acad. Sci. USA, 88:9578-9582 (1991)) as disclosed by Chevray and Nathans, Proc. Natl. Acad. Sci. USA, 89: 5789-5793 (1991). Many transcriptional activators, such as yeast GAL4, consist of two physically discrete modular domains, one acting as the DNA-binding domain, the other one functioning as the transcription-activation domain. The yeast expression system described in the foregoing publications (generally referred to as the "two-hybrid system") takes advantage of this property, and employs two hybrid proteins, one in which the target protein is fused to the DNA-binding domain of GAL4, and another, in which candidate activating proteins are fused to the activation domain. The expression of a GAL1-*lacZ* reporter gene under control of a GAL4-activated promoter depends on reconstitution of GAL4 activity via protein-protein interaction. Colonies containing interacting polypeptides are detected with a chromogenic substrate for β -galactosidase. A complete kit (MATCHMAKER™) for identifying protein-protein interactions between two specific proteins using the two-hybrid technique is commercially available from Clontech. This system can also be extended to map protein

domains involved in specific protein interactions as well as to pinpoint amino acid residues that are crucial for these interactions.

Compounds that interfere with the interaction of a gene encoding a PRO polypeptide identified herein and other intra- or extracellular components can be tested as follows: usually a reaction mixture is prepared containing the product of the gene and the intra- or extracellular component under conditions and for a time allowing for the interaction and binding of the two products. To test the ability of a candidate compound to inhibit binding, the reaction is run in the absence and in the presence of the test compound. In addition, a placebo may be added to a third reaction mixture, to serve as positive control. The binding (complex formation) between the test compound and the intra- or extracellular component present in the mixture is monitored as described hereinabove. The formation of a complex in the control reaction(s) but not in the reaction mixture containing the test compound indicates that the test compound interferes with the interaction of the test compound and its reaction partner.

To assay for antagonists, the PRO polypeptide may be added to a cell along with the compound to be screened for a particular activity and the ability of the compound to inhibit the activity of interest in the presence of the PRO polypeptide indicates that the compound is an antagonist to the PRO polypeptide. Alternatively, antagonists may be detected by combining the PRO polypeptide and a potential antagonist with membrane-bound PRO polypeptide receptors or recombinant receptors under appropriate conditions for a competitive inhibition assay. The PRO polypeptide can be labeled, such as by radioactivity, such that the number of PRO polypeptide molecules bound to the receptor can be used to determine the effectiveness of the potential antagonist. The gene encoding the receptor can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Coligan et al., Current Protocols in Immun., 1(2): Chapter 5 (1991). Preferably, expression cloning is employed wherein polyadenylated RNA is prepared from a cell responsive to the PRO polypeptide and a cDNA library created from this RNA is divided into pools and used to transfect COS cells or other cells that are not responsive to the PRO polypeptide. Transfected cells that are grown on glass slides are exposed to labeled PRO polypeptide. The PRO polypeptide can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase. Following fixation and incubation, the slides are subjected to autoradiographic analysis. Positive pools are identified and sub-pools are prepared and re-transfected using an interactive sub-pooling and re-screening process, eventually yielding a single clone that encodes the putative receptor.

As an alternative approach for receptor identification, labeled PRO polypeptide can be photoaffinity-linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is resolved by PAGE and exposed to X-ray film. The labeled complex containing the receptor can be excised, resolved into peptide fragments, and subjected to protein micro-sequencing. The amino acid sequence obtained from micro-sequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the gene encoding the putative receptor.

In another assay for antagonists, mammalian cells or a membrane preparation expressing the receptor would be incubated with labeled PRO polypeptide in the presence of the candidate compound. The ability of the compound to enhance or block this interaction could then be measured.

More specific examples of potential antagonists include an oligonucleotide that binds to the fusions of immunoglobulin with PRO polypeptide, and, in particular, antibodies including, without limitation, poly- and monoclonal antibodies and antibody fragments, single-chain antibodies, anti-idiotypic antibodies, and chimeric or humanized versions of such antibodies or fragments, as well as human antibodies and antibody fragments. Alternatively, a potential antagonist may be a closely related protein, for example, a mutated form of the PRO polypeptide that recognizes the receptor but imparts no effect, thereby competitively inhibiting the action of the PRO polypeptide.

Another potential PRO polypeptide antagonist is an antisense RNA or DNA construct prepared using antisense technology, where, e.g., an antisense RNA or DNA molecule acts to block directly the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes the mature PRO polypeptides herein, is used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res., 6:3073 (1979); Cooney et al., Science, 241: 456 (1988); Dervan et al., Science, 251:1360 (1991)), thereby preventing transcription and the production of the PRO polypeptide. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into the PRO polypeptide (antisense - Okano, Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression (CRC Press: Boca Raton, FL, 1988). The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of the PRO polypeptide. When antisense DNA is used, oligodeoxyribonucleotides derived from the translation-initiation site, e.g., between about -10 and +10 positions of the target gene nucleotide sequence, are preferred.

Potential antagonists include small molecules that bind to the active site, the receptor binding site, or growth factor or other relevant binding site of the PRO polypeptide, thereby blocking the normal biological activity of the PRO polypeptide. Examples of small molecules include, but are not limited to, small peptides or peptide-like molecules, preferably soluble peptides, and synthetic non-peptidyl organic or inorganic compounds.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. Ribozymes act by sequence-specific hybridization to the complementary target RNA, followed by endonucleolytic cleavage. Specific ribozyme cleavage sites within a potential RNA target can be identified by known techniques. For further details see, e.g., Rossi, Current Biology, 4:469-471 (1994), and PCT publication No. WO 97/33551 (published September 18, 1997).

Nucleic acid molecules in triple-helix formation used to inhibit transcription should be single-stranded and composed of deoxynucleotides. The base composition of these oligonucleotides is designed such that it promotes triple-helix formation via Hoogsteen base-pairing rules, which generally require sizeable stretches of purines or pyrimidines on one strand of a duplex. For further details see, e.g., PCT publication No. WO 97/33551, *supra*.

These small molecules can be identified by any one or more of the screening assays discussed hereinabove and/or by any other screening techniques well known for those skilled in the art.

Diagnostic and therapeutic uses of the herein disclosed molecules may also be based upon the positive functional assay hits disclosed and described below.

5 F. Anti-PRO Antibodies

The present invention further provides anti-PRO antibodies. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

1. Polyclonal Antibodies

10 The anti-PRO antibodies may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the PRO polypeptide or a fusion protein thereof. It may be useful to conjugate
15 the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

20 2. Monoclonal Antibodies

The anti-PRO antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an
25 immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

The immunizing agent will typically include the PRO polypeptide or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then
30 fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or
35 survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of

HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against PRO. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, supra]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., supra] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent

heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

In vitro methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

3. Human and Humanized Antibodies

The anti-PRO antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-329 (1988); and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeven et al., Science, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985) and

Boerner et al., J. Immunol., 147(1):86-95 (1991)]. Similarly, human antibodies can be made by introducing of human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks *et al.*, Bio/Technology 10, 779-783 (1992); Lonberg *et al.*, Nature 368 856-859 (1994); Morrison, Nature 368, 812-13 (1994); Fishwild *et al.*, Nature Biotechnology 14, 845-51 (1996); Neuberger, Nature Biotechnology 14, 826 (1996); Lonberg and Huszar, Intern. Rev. Immunol. 13 65-93 (1995).

The antibodies may also be affinity matured using known selection and/or mutagenesis methods as described above. Preferred affinity matured antibodies have an affinity which is five times, more preferably 10 times, even more preferably 20 or 30 times greater than the starting antibody (generally murine, humanized or human) from which the matured antibody is prepared.

4. Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for the PRO, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, Nature, 305:537-539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, EMBO J., 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh *et al.*, Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain.

In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

5 Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. $F(ab')_2$ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan *et al.*, Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate $F(ab')_2$ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

10 Fab' fragments may be directly recovered from *E. coli* and chemically coupled to form bispecific antibodies. Shalaby *et al.*, J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody $F(ab')_2$ molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling *in vitro* to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

20 Various technique for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny *et al.*, J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger *et al.*, Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber *et al.*, J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared.

35 Tutt *et al.*, J. Immunol. 147:60 (1991).

Exemplary bispecific antibodies may bind to two different epitopes on a given PRO polypeptide herein. Alternatively, an anti-PRO polypeptide arm may be combined with an arm which binds to a triggering molecule

on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular PRO polypeptide. Bispecific antibodies may also be used to localize cytotoxic agents to cells which express a particular PRO polypeptide. These antibodies possess a PRO-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the PRO polypeptide and further binds tissue factor (TF).

5. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

6. Effector Function Engineering

It may be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) may be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron *et al.*, J. Exp Med., **176**: 1191-1195 (1992) and Shopes, J. Immunol., **148**: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff *et al.* Cancer Research, **53**: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson *et al.*, Anti-Cancer Drug Design, **3**: 219-230 (1989).

7. Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolacca americana* proteins

(PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi , ^{131}I , ^{131}In , ^{90}Y , and ^{186}Re .

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al.*, Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody may be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (*e.g.*, avidin) that is conjugated to a cytotoxic agent (*e.g.*, a radionucleotide).

8. Immunoliposomes

The antibodies disclosed herein may also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein *et al.*, Proc. Natl. Acad. Sci. USA, 82: 3688 (1985); Hwang *et al.*, Proc. Natl. Acad. Sci. USA, 77: 4030 (1980); and U.S. Pat. Nos. 4,485,045 and 4,544,545. Liposomes with enhanced circulation time are disclosed in U.S. Patent No. 5,013,556.

Particularly useful liposomes can be generated by the reverse-phase evaporation method with a lipid composition comprising phosphatidylcholine, cholesterol, and PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of the antibody of the present invention can be conjugated to the liposomes as described in Martin *et al.*, J. Biol. Chem., 257: 286-288 (1982) via a disulfide-interchange reaction. A chemotherapeutic agent (such as Doxorubicin) is optionally contained within the liposome. See Gabizon *et al.*, J. National Cancer Inst., 81(19): 1484 (1989).

9. Pharmaceutical Compositions of Antibodies

Antibodies specifically binding a PRO polypeptide identified herein, as well as other molecules identified by the screening assays disclosed hereinbefore, can be administered for the treatment of various disorders in the form of pharmaceutical compositions.

If the PRO polypeptide is intracellular and whole antibodies are used as inhibitors, internalizing antibodies are preferred. However, lipofections or liposomes can also be used to deliver the antibody, or an antibody fragment, into cells. Where antibody fragments are used, the smallest inhibitory fragment that

specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable-region sequences of an antibody, peptide molecules can be designed that retain the ability to bind the target protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology. See, *e.g.*, Marasco *et al.*, Proc. Natl. Acad. Sci. USA, 90: 7889-7893 (1993). The formulation herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition may comprise an agent that enhances its function, such as, for example, a cytotoxic agent, cytokine, chemotherapeutic agent, or growth-inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

The active ingredients may also be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacrylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles, and nanocapsules) or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences, *supra*.

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

Sustained-release preparations may be prepared. Suitable examples of sustained-release preparations include semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, *e.g.*, films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and γ ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOTTM (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid. While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods. When encapsulated antibodies remain in the body for a long time, they may denature or aggregate as a result of exposure to moisture at 37°C, resulting in a loss of biological activity and possible changes in immunogenicity. Rational strategies can be devised for stabilization depending on the mechanism involved. For example, if the aggregation mechanism is discovered to be intermolecular S-S bond formation through thio-disulfide interchange, stabilization may be achieved by modifying sulfhydryl residues, lyophilizing from acidic solutions, controlling moisture content, using appropriate additives, and developing specific polymer matrix compositions.

G. Uses for anti-PRO Antibodies

The anti-PRO antibodies of the invention have various utilities. For example, anti-PRO antibodies may be used in diagnostic assays for PRO, *e.g.*, detecting its expression (and in some cases, differential expression) in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC

Press, Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as ^3H , ^{14}C , ^{32}P , ^{35}S , or ^{125}I , a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., *Nature*, 144:945 (1962); David et al., *Biochemistry*, 13:1014 (1974); Pain et al., *J. Immunol. Meth.*, 40:219 (1981); and Nygren, *J. Histochem. and Cytochem.*, 30:407 (1982).

Anti-PRO antibodies also are useful for the affinity purification of PRO from recombinant cell culture or natural sources. In this process, the antibodies against PRO are immobilized on a suitable support, such as Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO from the antibody.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Manassas, VA.

EXAMPLE 1: Extracellular Domain Homology Screening to Identify Novel Polypeptides and cDNA Encoding Therefor

The extracellular domain (ECD) sequences (including the secretion signal sequence, if any) from about 950 known secreted proteins from the Swiss-Prot public database were used to search EST databases. The EST databases included public databases (e.g., Dayhoff, GenBank), and proprietary databases (e.g. LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST-2 (Altschul et al., *Methods in Enzymology* 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequences. Those comparisons with a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, WA).

Using this extracellular domain homology screen, consensus DNA sequences were assembled relative to the other identified EST sequences using phrap. In addition, the consensus DNA sequences obtained were

often (but not always) extended using repeated cycles of BLAST or BLAST-2 and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above.

Based upon the consensus sequences obtained as described above, oligonucleotides were then synthesized and used to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for a PRO polypeptide. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

EXAMPLE 2: Isolation of cDNA clones by Amylase Screening

1. Preparation of oligo dT primed cDNA library

mRNA was isolated from a human tissue of interest using reagents and protocols from Invitrogen, San Diego, CA (Fast Track 2). This RNA was used to generate an oligo dT primed cDNA library in the vector pRK5D using reagents and protocols from Life Technologies, Gaithersburg, MD (Super Script Plasmid System). In this procedure, the double stranded cDNA was sized to greater than 1000 bp and the SalI/NotI linked cDNA was cloned into XhoI/NotI cleaved vector. pRK5D is a cloning vector that has an sp6 transcription initiation site followed by an SfiI restriction enzyme site preceding the XhoI/NotI cDNA cloning sites.

2. Preparation of random primed cDNA library

A secondary cDNA library was generated in order to preferentially represent the 5' ends of the primary cDNA clones. Sp6 RNA was generated from the primary library (described above), and this RNA was used to generate a random primed cDNA library in the vector pSST-AMY.0 using reagents and protocols from Life Technologies (Super Script Plasmid System, referenced above). In this procedure the double stranded cDNA was sized to 500-1000 bp, linked with blunt to NotI adaptors, cleaved with SfiI, and cloned into SfiI/NotI cleaved vector. pSST-AMY.0 is a cloning vector that has a yeast alcohol dehydrogenase promoter preceding the cDNA cloning sites and the mouse amylase sequence (the mature sequence without the secretion signal) followed by the yeast alcohol dehydrogenase terminator, after the cloning sites. Thus, cDNAs cloned into this vector that are fused in frame with amylase sequence will lead to the secretion of amylase from appropriately transfected yeast colonies.

3. Transformation and Detection

DNA from the library described in paragraph 2 above was chilled on ice to which was added electrocompetent DH10B bacteria (Life Technologies, 20 ml). The bacteria and vector mixture was then electroporated as recommended by the manufacturer. Subsequently, SOC media (Life Technologies, 1 ml) was added and the mixture was incubated at 37°C for 30 minutes. The transformants were then plated onto 20 standard 150 mm LB plates containing ampicillin and incubated for 16 hours (37°C). Positive colonies were scraped off the plates and the DNA was isolated from the bacterial pellet using standard protocols, e.g. CsCl-gradient. The purified DNA was then carried on to the yeast protocols below.

The yeast methods were divided into three categories: (1) Transformation of yeast with the plasmid/cDNA combined vector; (2) Detection and isolation of yeast clones secreting amylase; and (3) PCR amplification of the insert directly from the yeast colony and purification of the DNA for sequencing and further analysis.

The yeast strain used was HD56-5A (ATCC-90785). This strain has the following genotype: MAT alpha, ura3-52, leu2-3, leu2-112, his3-11, his3-15, MAL⁺, SUC⁺, GAL⁺. Preferably, yeast mutants can be employed that have deficient post-translational pathways. Such mutants may have translocation deficient alleles in *sec71*, *sec72*, *sec62*, with truncated *sec71* being most preferred. Alternatively, antagonists (including antisense nucleotides and/or ligands) which interfere with the normal operation of these genes, other proteins implicated in this post translation pathway (e.g., SEC61p, SEC72p, SEC62p, SEC63p, TDJ1p or SSA1p-4p) or the complex formation of these proteins may also be preferably employed in combination with the amylase-expressing yeast.

Transformation was performed based on the protocol outlined by Gietz et al., Nucl. Acid. Res., 20:1425 (1992). Transformed cells were then inoculated from agar into YEPD complex media broth (100 ml) and grown overnight at 30°C. The YEPD broth was prepared as described in Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 207 (1994). The overnight culture was then diluted to about 2×10^6 cells/ml (approx. OD₆₀₀=0.1) into fresh YEPD broth (500 ml) and regrown to 1×10^7 cells/ml (approx. OD₆₀₀=0.4-0.5).

The cells were then harvested and prepared for transformation by transfer into GS3 rotor bottles in a Sorval GS3 rotor at 5,000 rpm for 5 minutes, the supernatant discarded, and then resuspended into sterile water, and centrifuged again in 50 ml falcon tubes at 3,500 rpm in a Beckman GS-6KR centrifuge. The supernatant was discarded and the cells were subsequently washed with LiAc/TE (10 ml, 10 mM Tris-HCl, 1 mM EDTA pH 7.5, 100 mM Li₂OOCCH₃), and resuspended into LiAc/TE (2.5 ml).

Transformation took place by mixing the prepared cells (100 µl) with freshly denatured single stranded salmon testes DNA (Lofstrand Labs, Gaithersburg, MD) and transforming DNA (1 µg, vol. < 10 µl) in microfuge tubes. The mixture was mixed briefly by vortexing, then 40% PEG/TE (600 µl, 40% polyethylene glycol-4000, 10 mM Tris-HCl, 1 mM EDTA, 100 mM Li₂OOCCH₃, pH 7.5) was added. This mixture was gently mixed and incubated at 30°C while agitating for 30 minutes. The cells were then heat shocked at 42°C for 15 minutes, and the reaction vessel centrifuged in a microfuge at 12,000 rpm for 5-10 seconds, decanted and resuspended into TE (500 µl, 10 mM Tris-HCl, 1 mM EDTA pH 7.5) followed by recentrifugation. The cells

were then diluted into TE (1 ml) and aliquots (200 μ l) were spread onto the selective media previously prepared in 150 mm growth plates (VWR).

Alternatively, instead of multiple small reactions, the transformation was performed using a single, large scale reaction, wherein reagent amounts were scaled up accordingly.

The selective media used was a synthetic complete dextrose agar lacking uracil (SCD-Ura) prepared as described in Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 208-210 (1994). Transformants were grown at 30°C for 2-3 days.

The detection of colonies secreting amylase was performed by including red starch in the selective growth media. Starch was coupled to the red dye (Reactive Red-120, Sigma) as per the procedure described by Biely et al., Anal. Biochem., 172:176-179 (1988). The coupled starch was incorporated into the SCD-Ura agar plates at a final concentration of 0.15% (w/v), and was buffered with potassium phosphate to a pH of 7.0 (50-100 mM final concentration).

The positive colonies were picked and streaked across fresh selective media (onto 150 mm plates) in order to obtain well isolated and identifiable single colonies. Well isolated single colonies positive for amylase secretion were detected by direct incorporation of red starch into buffered SCD-Ura agar. Positive colonies were determined by their ability to break down starch resulting in a clear halo around the positive colony visualized directly.

4. Isolation of DNA by PCR Amplification

When a positive colony was isolated, a portion of it was picked by a toothpick and diluted into sterile water (30 μ l) in a 96 well plate. At this time, the positive colonies were either frozen and stored for subsequent analysis or immediately amplified. An aliquot of cells (5 μ l) was used as a template for the PCR reaction in a 25 μ l volume containing: 0.5 μ l KlenTaq (Clontech, Palo Alto, CA); 4.0 μ l 10 mM dNTP's (Perkin Elmer-Cetus); 2.5 μ l Kentaq buffer (Clontech); 0.25 μ l forward oligo 1; 0.25 μ l reverse oligo 2; 12.5 μ l distilled water. The sequence of the forward oligonucleotide 1 was:

5'-TGTAACGACGGCCAGTTAAATAGACCTGCAATTATTAATCT-3' (SEQ ID NO:553)

The sequence of reverse oligonucleotide 2 was:

5'-CAGGAAACAGCTATGACCACCTGCACACCTGCAAATCCATT-3' (SEQ ID NO:554)

PCR was then performed as follows:

a.	Denature	92°C, 5 minutes
b.	3 cycles of:	
	Denature	92°C, 30 seconds
	Anneal	59°C, 30 seconds
	Extend	72°C, 60 seconds
c.	3 cycles of:	
	Denature	92°C, 30 seconds
	Anneal	57°C, 30 seconds
	Extend	72°C, 60 seconds
d.	25 cycles of:	
	Denature	92°C, 30 seconds
	Anneal	55°C, 30 seconds
	Extend	72°C, 60 seconds

e. Hold 4°C

The underlined regions of the oligonucleotides annealed to the ADH promoter region and the amylase region, respectively, and amplified a 307 bp region from vector pSST-AMY.0 when no insert was present. Typically, the first 18 nucleotides of the 5' end of these oligonucleotides contained annealing sites for the sequencing primers. Thus, the total product of the PCR reaction from an empty vector was 343 bp. However, signal sequence-fused cDNA resulted in considerably longer nucleotide sequences.

Following the PCR, an aliquot of the reaction (5 µl) was examined by agarose gel electrophoresis in a 1% agarose gel using a Tris-Borate-EDTA (TBE) buffering system as described by Sambrook et al., *supra*. Clones resulting in a single strong PCR product larger than 400 bp were further analyzed by DNA sequencing after purification with a 96 Qiaquick PCR clean-up column (Qiagen Inc., Chatsworth, CA).

EXAMPLE 3: Isolation of cDNA Clones Using Signal Algorithm Analysis

Various polypeptide-encoding nucleic acid sequences were identified by applying a proprietary signal sequence finding algorithm developed by Genentech, Inc. (South San Francisco, CA) upon ESTs as well as clustered and assembled EST fragments from public (e.g., GenBank) and/or private (LIFESEQ®, Incyte Pharmaceuticals, Inc., Palo Alto, CA) databases. The signal sequence algorithm computes a secretion signal score based on the character of the DNA nucleotides surrounding the first and optionally the second methionine codon(s) (ATG) at the 5'-end of the sequence or sequence fragment under consideration. The nucleotides following the first ATG must code for at least 35 unambiguous amino acids without any stop codons. If the first ATG has the required amino acids, the second is not examined. If neither meets the requirement, the candidate sequence is not scored. In order to determine whether the EST sequence contains an authentic signal sequence, the DNA and corresponding amino acid sequences surrounding the ATG codon are scored using a set of seven sensors (evaluation parameters) known to be associated with secretion signals. Use of this algorithm resulted in the identification of numerous polypeptide-encoding nucleic acid sequences.

EXAMPLE 4: Isolation of cDNA clones Encoding Human PRO Polypeptides

Using the techniques described in Examples 1 to 3 above, numerous full-length cDNA clones were identified as encoding PRO polypeptides as disclosed herein. These cDNAs were then deposited under the terms of the Budapest Treaty with the American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209, USA (ATCC) as shown in Table 7 below.

Table 7

<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
DNA16438-1387	209771	April 14, 1998
DNA19360-2552	203654	February 9, 1999
DNA33455-1548	PTA-127	May 25, 1999
DNA37155-2651	PTA-429	July 27, 1999
DNA38269-2654	PTA-432	July 27, 1999
DNA40619-1220	209525	December 10, 1997

Table 7 (cont')

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
	DNA44174-2513	203577	January 12, 1999
	DNA44675-2662	PTA-430	July 27, 1999
	DNA45408-2615	PTA-203	June 8, 1999
5	DNA48606-1479	203040	July 1, 1998
	DNA52753-2656	PTA-611	August 31, 1999
	DNA53915-1258	209593	January 21, 1998
	DNA53991-2553	203649	February 9, 1999
	DNA54009-2517	203574	January 12, 1999
10	DNA56055-1643	PTA-129	May 25, 1999
	DNA57033-1403	209905	May 27, 1998
	DNA57252-1453	203585	January 12, 1999
	DNA58799-1652	203665	February 9, 1999
	DNA59770-2652	PTA-427	July 27, 1999
15	DNA59774-2665	PTA-615	August 31, 1999
	DNA60281-2518	203582	January 12, 1999
	DNA60736-2559	203838	March 9, 1999
	DNA61875-2653	PTA-428	July 27, 1999
	DNA62312-2558	203836	March 9, 1999
20	DNA62849-1604	PTA-205	June 8, 1999
	DNA66307-2661	PTA-431	July 27, 1999
	DNA66677-2535	203659	February 9, 1999
	DNA71235-1706	203584	January 12, 1999
	DNA71289-2547	PTA-126	May 25, 1999
25	DNA73775-1707	PTA-128	May 25, 1999
	DNA76385-1692	203664	February 9, 1999
	DNA76395-2527	203578	January 12, 1999
	DNA77622-2516	203554	December 22, 1998
	DNA77629-2573	203850	March 16, 1999
30	DNA77645-2648	PTA-45	May 11, 1999
	DNA79302-2521	203545	December 22, 1998
	DNA79865-2519	203544	December 22, 1998
	DNA80135-2655	PTA-234	June 15, 1999
	DNA80794-2568	203848	March 16, 1999
35	DNA80796-2523	203555	December 22, 1998
	DNA80840-2605	203949	April 20, 1999
	DNA80899-2501	203539	December 15, 1998
	DNA81228-2580	203871	March 23, 1999
	DNA81761-2583	203862	March 23, 1999
40	DNA82358-2738	PTA-510	August 10, 1999
	DNA82364-2538	203603	January 20, 1999
	DNA82424-2566	203813	March 2, 1999
	DNA82430-2557	203812	March 2, 1999
	DNA83500-2506	203391	October 29, 1998
45	DNA83509-2612	203965	April 27, 1999
	DNA83560-2569	203816	March 2, 1999
	DNA84139-2555	203814	March 2, 1999
	DNA84141-2556	203810	March 2, 1999
	DNA84142-2613	PTA-22	May 4, 1999
50	DNA84318-2520	203580	January 12, 1999
	DNA84909-2590	203889	March 30, 1999
	DNA84912-2610	203964	April 27, 1999
	DNA84925-2514	203548	December 22, 1998
	DNA84928-2564	203817	March 2, 1999
55	DNA84932-2657	PTA-235	June 15, 1999

Table 7 (cont')

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
	DNA86592-2607	203968	April 27, 1999
	DNA86594-2587	203894	March 30, 1999
	DNA86647-2591	203893	March 30, 1999
5	DNA87185-2563	203811	March 2, 1999
	DNA87656-2582	203867	March 23, 1999
	DNA87974-2609	203963	April 27, 1999
	DNA88001-2565	203815	March 2, 1999
	DNA88004-2575	203890	March 30, 1999
10	DNA89220-2608	PTA-130	May 25, 1999
	DNA89947-2618	203970	April 27, 1999
	DNA90842-2574	203845	March 16, 1999
	DNA91775-2581	203861	March 23, 1999
	DNA91779-2571	203844	March 16, 1999
15	DNA92217-2697	PTA-513	August 10, 1999
	DNA92219-2541	203663	February 9, 1999
	DNA92223-2567	203851	March 16, 1999
	DNA92225-2603	203950	April 20, 1999
	DNA92232-2589	203895	March 30, 1999
20	DNA92233-2599	PTA-134	May 25, 1999
	DNA92243-2549	203852	March 16, 1999
	DNA92253-2671	PTA-258	June 22, 1999
	DNA92254-2672	PTA-259	June 22, 1999
	DNA92255-2584	203866	March 23, 1999
25	DNA92269-2570	203853	March 16, 1999
	DNA92288-2588	203892	March 30, 1999
	DNA92290-2550	203847	March 16, 1999
	DNA93012-2622	PTA-21	May 4, 1999
	DNA93020-2642	PTA-121	May 25, 1999
30	DNA94830-2604	203951	April 20, 1999
	DNA94833-2579	203869	March 23, 1999
	DNA94838-2658	PTA-232	June 15, 1999
	DNA94844-2686	PTA-385	July 20, 1999
	DNA94854-2586	203864	March 23, 1999
35	DNA96868-2677	PTA-262	June 22, 1999
	DNA96871-2683	PTA-381	July 20, 1999
	DNA96880-2624	PTA-15	May 4, 1999
	DNA96986-2660	PTA-239	June 15, 1999
	DNA96988-2685	PTA-384	July 20, 1999
40	DNA96995-2709	PTA-475	August 3, 1999
	DNA97004-2562	203854	March 16, 1999
	DNA97005-2687	PTA-378	July 20, 1999
	DNA97009-2668	PTA-257	June 22, 1999
	DNA97013-2667	PTA-231	June 15, 1999
45	DNA98380-2690	PTA-388	July 20, 1999
	DNA98561-2696	PTA-620	August 31, 1999
	DNA98575-2644	PTA-118	May 25, 1999
	DNA98593-2694	PTA-477	August 3, 1999
	DNA98600-2703	PTA-488	August 3, 1999
50	DNA99391-2572	203849	March 16, 1999
	DNA99393-2560	203837	March 9, 1999
	DNA100276-2684	PTA-380	July 20, 1999
	DNA100312-2645	PTA-44	May 11, 1999
	DNA100902-2646	PTA-42	May 11, 1999
55	DNA102899-2679	PTA-123	May 25, 1999

Table 7 (cont')

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
	DNA104875-2720	PTA-482	August 3, 1999
	DNA105680-2710	PTA-483	August 3, 1999
	DNA105779-2708	PTA-485	August 3, 1999
5	DNA105794-2695	PTA-480	August 3, 1999
	DNA105838-2702	PTA-476	August 3, 1999
	DNA107698-2715	PTA-472	August 3, 1999
	DNA107701-2711	PTA-487	August 3, 1999
	DNA107781-2707	PTA-484	August 3, 1999
10	DNA108670-2744	PTA-546	August 17, 1999
	DNA108688-2725	PTA-515	August 10, 1999
	DNA108769-2765	PTA-861	October 19, 1999
	DNA108935-2721	PTA-518	August 10, 1999
	DNA110700-2716	PTA-512	August 10, 1999
15	DNA111750-2706	PTA-489	August 3, 1999
	DNA123430-2755	PTA-614	August 31, 1999
	DNA125154-2785	PTA-957	November 16, 1999
	DNA142238-2768	PTA-819	October 5, 1999
	DNA22779-1130	209280	September 18, 1997
20	DNA26847-1395	209772	April 14, 1998
	DNA27864-1155	209375	October 16, 1997
	DNA27865-1091	209296	September 23, 1997
	DNA28497-1130	209279	September 18, 1997
	DNA29101-1122	209653	March 5, 1998
25	DNA32286-1191	209385	October 16, 1997
	DNA32288-1132	209261	September 16, 1997
	DNA32290-1164	209384	October 16, 1997
	DNA32292-1131	209258	September 16, 1997
	DNA32298-1132	209257	September 16, 1997
30	DNA33085-1110	209087	May 30, 1997
	DNA33087-1158	209381	October 16, 1997
	DNA33089-1132	209262	September 16, 1997
	DNA33092-1202	209420	October 28, 1997
	DNA33094-1131	209256	September 16, 1997
35	DNA33107-1135	209251	September 16, 1997
	DNA33221-1133	209263	September 16, 1997
	DNA33223-1136	209264	September 16, 1997
	DNA33460-1166	209376	October 16, 1997
	DNA33473-1176	209391	October 17, 1997
40	DNA33785-1143	209417	October 28, 1997
	DNA33786-1132	209253	September 16, 1997
	DNA34353-1428	209855	May 12, 1998
	DNA34392-1170	209526	December 10, 1997
	DNA34434-1139	209252	September 16, 1997
45	DNA35558-1167	209374	October 16, 1997
	DNA35595-1228	209528	December 10, 1997
	DNA35638-1216	209265	September 16, 1997
	DNA35639-1172	209396	October 17, 1997
	DNA35663-1129	209201	August 18, 1997
50	DNA35674-1142	209416	October 28, 1997
	DNA35841-1173	209403	October 17, 1997
	DNA35916-1161	209419	October 28, 1997
	DNA35918-1174	209402	October 17, 1997
	DNA36350-1158	209378	October 16, 1997
55	DNA37140-1234	209489	November 21, 1997

Table 7 (cont')

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
	DNA37150-1178	209401	October 17, 1997
	DNA38260-1180	209397	October 17, 1997
	DNA40021-1154	209389	October 17, 1997
5	DNA40587-1231	209438	November 7, 1997
	DNA40592-1242	209492	November 21, 1997
	DNA40620-1183	209388	October 17, 1997
	DNA40628-1216	209432	November 7, 1997
	DNA40981-1234	209439	November 7, 1997
10	DNA40982-1235	209433	November 7, 1997
	DNA41234-1242	209618	February 5, 1998
	DNA43046-1225	209484	November 21, 1997
	DNA43316-1237	209487	November 21, 1997
	DNA44167-1243	209434	November 7, 1997
15	DNA44184-1319	209704	March 26, 1998
	DNA44194-1317	209808	April 28, 1998
	DNA44196-1353	209847	May 6, 1998
	DNA45419-1252	209616	February 5, 1998
	DNA46777-1253	209619	February 5, 1998
20	DNA47394-1572	203109	August 11, 1998
	DNA48331-1329	209715	March 31, 1998
	DNA48336-1309	209669	March 11, 1998
	DNA49142-1430	203002	June 23, 1998
	DNA49646-1327	209705	March 26, 1998
25	DNA49821-1562	209981	June 16, 1998
	DNA49829-1346	209749	April 7, 1998
	DNA50921-1458	209859	May 12, 1998
	DNA52187-1354	209845	May 6, 1998
	DNA52196-1348	209748	April 7, 1998
30	DNA52598-1518	203107	August 11, 1998
	DNA54228-1366	209801	April 23, 1998
	DNA56047-1456	209948	June 9, 1998
	DNA56112-1379	209883	May 20, 1998
	DNA56113-1378	203049	July 1, 1998
35	DNA56352-1358	209846	May 6, 1998
	DNA56433-1406	209857	May 12, 1998
	DNA56439-1376	209864	May 14, 1998
	DNA57530-1375	209880	May 20, 1998
	DNA57689-1385	209869	May 14, 1998
40	DNA57690-1374	209950	June 9, 1998
	DNA57693-1424	203008	June 23, 1998
	DNA57838-1337	203014	June 23, 1998
	DNA58721-1475	203110	August 11, 1998
	DNA59205-1421	203009	June 23, 1998
45	DNA59215-1425	209961	June 9, 1998
	DNA59220-1514	209962	June 9, 1998
	DNA59294-1381	209866	May 14, 1998
	DNA59488-1603	203157	August 25, 1998
	DNA59588-1571	203106	August 11, 1998
50	DNA59606-1471	209945	June 9, 1998
	DNA59620-1463	209989	June 16, 1998
	DNA59767-1489	203108	August 11, 1998
	DNA59777-1480	203111	August 11, 1998
	DNA59814-1486	203359	October 20, 1998
55	DNA59839-1461	209988	June 16, 1998

Table 7 (cont')

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
	DNA59846-1503	209978	June 16, 1998
	DNA59847-1511	203098	August 4, 1998
	DNA60615-1483	209980	June 16, 1998
5	DNA60621-1516	203091	August 4, 1998
	DNA60622-1525	203090	August 4, 1998
	DNA60627-1508	203092	August 4, 1998
	DNA60764-1533	203452	November 10, 1998
	DNA60775-1532	203173	September 1, 1998
10	DNA61185-1646	203464	November 17, 1998
	DNA61873-1574	203132	August 18, 1998
	DNA62306-1570	203254	September 9, 1998
	DNA62808-1582	203358	October 20, 1998
	DNA62814-1521	203093	August 4, 1998
15	DNA64885-1529	203457	November 3, 1998
	DNA64886-1601	203241	September 9, 1998
	DNA64888-1542	203249	September 9, 1998
	DNA64889-1541	203250	September 9, 1998
	DNA64890-1612	203131	August 18, 1998
20	DNA64903-1553	203223	September 15, 1998
	DNA64905-1558	203233	September 15, 1998
	DNA65402-1540	203252	September 9, 1998
	DNA65405-1547	203476	November 17, 1998
	DNA65412-1523	203094	August 4, 1998
25	DNA66309-1538	203235	September 15, 1998
	DNA66667-1596	203267	September 22, 1998
	DNA66675-1587	203282	September 22, 1998
	DNA68818-2536	203657	February 9, 1999
	DNA68864-1629	203276	September 22, 1998
30	DNA68872-1620	203160	August 25, 1998
	DNA71159-1617	203135	August 18, 1998
	DNA73727-1673	203459	November 3, 1998
	DNA73739-1645	203270	September 22, 1998
	DNA76400-2528	203573	January 12, 1999
35	DNA76510-2504	203477	November 17, 1998
	DNA76529-1666	203315	October 6, 1998
	DNA76538-1670	203313	October 6, 1998
	DNA77301-1708	203407	October 27, 1998
	DNA77624-2515	203553	December 22, 1998
40	DNA79230-2525	203549	December 22, 1998
	DNA79862-2522	203550	December 22, 1998
	DNA80145-2594	PTA-204	June 8, 1999
	DNA83500-2506	203391	October 29, 1998
	DNA84917-2597	203863	March 23, 1999
45	DNA92218-2554	203834	March 9, 1999
	DNA96042-2682	PTA-382	July 20, 1999

These deposits were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposits will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of

the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC § 122 and the Commissioner's rules pursuant thereto (including 37 CFR § 1.14 with particular reference to 886 OG 638).

5 The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

10 EXAMPLE 5: Use of PRO as a hybridization probe

 The following method describes use of a nucleotide sequence encoding PRO as a hybridization probe.

 DNA comprising the coding sequence of full-length or mature PRO as disclosed herein is employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO) in human
15 tissue cDNA libraries or human tissue genomic libraries.

 Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled PRO-derived probe to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed
20 in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

 DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO can then be identified using standard techniques known in the art.

EXAMPLE 6: Expression of PRO in *E. coli*

25 This example illustrates preparation of an unglycosylated form of PRO by recombinant expression in *E. coli*.

 The DNA sequence encoding PRO is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is
30 pBR322 (derived from *E. coli*; see Bolivar et al., Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the PRO coding region, lambda transcriptional terminator, and an
35 argU gene.

 The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., supra. Transformants are identified by their ability to grow on LB plates and antibiotic resistant

colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

5 After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO protein can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

10 PRO may be expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding PRO is initially amplified using selected PCR primers. The primers will contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences are then ligated into an expression vector, which is used to transform an *E. coli* host based on strain 52 (W3110
15 fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq). Transformants are first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an O.D.600 of 3-5 is reached. Cultures are then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g (NH₄)₂SO₄, 0.71 g sodium citrate•2H₂O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 mL water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM MgSO₄) and grown for approximately 20-30 hours at 30°C with shaking. Samples are
20 removed to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets are frozen until purification and refolding.

E. coli paste from 0.5 to 1 L fermentations (6-10 g pellets) is resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution is stirred overnight at 4°C. This step results
25 in a denatured protein with all cysteine residues blocked by sulfitolization. The solution is centrifuged at 40,000 rpm in a Beckman Ultracentrifuge for 30 min. The supernatant is diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. The clarified extract is loaded onto a 5 ml Qiagen Ni-NTA metal chelate column equilibrated in the metal chelate column buffer. The column is washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein is eluted with buffer containing 250 mM imidazole. Fractions containing the
30 desired protein are pooled and stored at 4°C. Protein concentration is estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

The proteins are refolded by diluting the sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM EDTA.
35 Refolding volumes are chosen so that the final protein concentration is between 50 to 100 micrograms/ml. The refolding solution is stirred gently at 4°C for 12-36 hours. The refolding reaction is quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the

solution is filtered through a 0.22 micron filter and acetonitrile is added to 2-10% final concentration. The refolded protein is chromatographed on a Poros R1/H reversed phase column using a mobile buffer of 0.1% TFA with elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A280 absorbance are analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein are pooled. Generally, the properly refolded species of most proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded forms of proteins from the desired form, the reversed phase step also removes endotoxin from the samples.

Fractions containing the desired folded PRO polypeptide are pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins are formulated into 20 mM Hepes, pH 6.8 with 0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 7: Expression of PRO in mammalian cells

This example illustrates preparation of a potentially glycosylated form of PRO by recombinant expression in mammalian cells.

The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO DNA using ligation methods such as described in Sambrook et al., *supra*. The resulting vector is called pRK5-PRO.

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 µg pRK5-PRO DNA is mixed with about 1 µg DNA encoding the VA RNA gene [Thimmappaya et al., *Cell*, 31:543 (1982)] and dissolved in 500 µl of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M CaCl₂. To this mixture is added, dropwise, 500 µl of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM NaPO₄, and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 µCi/ml ³⁵S-cysteine and 200 µCi/ml ³⁵S-methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of PRO polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO may be introduced into 293 cells transiently using the dextran sulfate method described by Sompanyrac et al., Proc. Natl. Acad. Sci., 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 µg pRK5-PRO DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 µg/ml bovine insulin and 0.1 µg/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, PRO can be expressed in CHO cells. The pRK5-PRO can be transfected into CHO cells using known reagents such as CaPO₄ or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ³⁵S-methionine. After determining the presence of PRO polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO can then be concentrated and purified by any selected method.

Epitope-tagged PRO may also be expressed in host CHO cells. The PRO may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO can then be concentrated and purified by any selected method, such as by Ni²⁺-chelate affinity chromatography.

PRO may also be expressed in CHO and/or COS cells by a transient expression procedure or in CHO cells by another stable expression procedure.

Stable expression in CHO cells is performed using the following procedure. The proteins are expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g. extracellular domains) of the respective proteins are fused to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or is a poly-His tagged form.

Following PCR amplification, the respective DNAs are subcloned in a CHO expression vector using standard techniques as described in Ausubel et al., Current Protocols of Molecular Biology, Unit 3.16, John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used expression in CHO cells is as described in Lucas et al., Nucl. Acids Res. 24:9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

Twelve micrograms of the desired plasmid DNA is introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect[®] (Quiagen), Dosper[®] or Eugene[®] (Boehringer

Mannheim). The cells are grown as described in Lucas et al., *supra*. Approximately 3×10^7 cells are frozen in an ampule for further growth and production as described below.

The ampules containing the plasmid DNA are thawed by placement into water bath and mixed by vortexing. The contents are pipetted into a centrifuge tube containing 10 mLs of media and centrifuged at 1000 rpm for 5 minutes. The supernatant is aspirated and the cells are resuspended in 10 mL of selective media (0.2 μ m filtered PS20 with 5% 0.2 μ m diafiltered fetal bovine serum). The cells are then aliquoted into a 100 mL spinner containing 90 mL of selective media. After 1-2 days, the cells are transferred into a 250 mL spinner filled with 150 mL selective growth medium and incubated at 37°C. After another 2-3 days, 250 mL, 500 mL and 2000 mL spinners are seeded with 3×10^5 cells/mL. The cell media is exchanged with fresh media by centrifugation and resuspension in production medium. Although any suitable CHO media may be employed, a production medium described in U.S. Patent No. 5,122,469, issued June 16, 1992 may actually be used. A 3L production spinner is seeded at 1.2×10^6 cells/mL. On day 0, the cell number pH is determined. On day 1, the spinner is sampled and sparging with filtered air is commenced. On day 2, the spinner is sampled, the temperature shifted to 33°C, and 30 mL of 500 g/L glucose and 0.6 mL of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning 365 Medical Grade Emulsion) taken. Throughout the production, the pH is adjusted as necessary to keep it at around 7.2. After 10 days, or until the viability dropped below 70%, the cell culture is harvested by centrifugation and filtering through a 0.22 μ m filter. The filtrate was either stored at 4°C or immediately loaded onto columns for purification.

For the poly-His tagged constructs, the proteins are purified using a Ni-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column is washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein is subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc-containing) constructs are purified from the conditioned media as follows. The conditioned medium is pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 ml fractions into tubes containing 275 μ L of 1 M Tris buffer, pH 9. The highly purified protein is subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity is assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 8: Expression of PRO in Yeast

The following method describes recombinant expression of PRO in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO from the ADH2/GAPDH promoter. DNA encoding PRO and the promoter is inserted into suitable restriction enzyme

sites in the selected plasmid to direct intracellular expression of PRO. For secretion, DNA encoding PRO can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, a native PRO signal peptide or other mammalian signal peptide, or, for example, a yeast alpha-factor or invertase secretory signal/leader sequence, and linker sequences (if needed) for expression of PRO.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant PRO can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing PRO may further be purified using selected column chromatography resins.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 9: Expression of PRO in Baculovirus-Infected Insect Cells

The following method describes recombinant expression of PRO in Baculovirus-infected insect cells.

The sequence coding for PRO is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the sequence encoding PRO or the desired portion of the coding sequence of PRO such as the sequence encoding the extracellular domain of a transmembrane protein or the sequence encoding the mature protein if the protein is extracellular is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 - 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression are performed as described by O'Reilley et al., Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

Expressed poly-his tagged PRO can then be purified, for example, by Ni²⁺-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., Nature, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM MgCl₂; 0.1 mM EDTA; 10% glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% glycerol, pH 7.8) and filtered through a 0.45 μm filter. A Ni²⁺-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A₂₈₀ with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM

phosphate; 300 mM NaCl, 10% glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A_{280} baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or Western blot with Ni^{2+} -NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His₁₀-tagged PRO are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 10: Preparation of Antibodies that Bind PRO

This example illustrates preparation of monoclonal antibodies which can specifically bind PRO.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, *supra*. Immunogens that may be employed include purified PRO, fusion proteins containing PRO, and cells expressing recombinant PRO on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against PRO. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against PRO is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

EXAMPLE 11: Purification of PRO Polypeptides Using Specific Antibodies

Native or recombinant PRO polypeptides may be purified by a variety of standard techniques in the art of protein purification. For example, pro-PRO polypeptide, mature PRO polypeptide, or pre-PRO polypeptide is purified by immunoaffinity chromatography using antibodies specific for the PRO polypeptide of interest. In general, an immunoaffinity column is constructed by covalently coupling the anti-PRO polypeptide antibody to an activated chromatographic resin.

Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSE™ (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

Such an immunoaffinity column is utilized in the purification of PRO polypeptide by preparing a fraction from cells containing PRO polypeptide in a soluble form. This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by other methods well known in the art. Alternatively, soluble PRO polypeptide containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

A soluble PRO polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PRO polypeptide (e.g., high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/PRO polypeptide binding (e.g., a low pH buffer such as approximately pH 2-3, or a high concentration of a chaotrope such as urea or thiocyanate ion), and PRO polypeptide is collected.

EXAMPLE 12: Drug Screening

This invention is particularly useful for screening compounds by using PRO polypeptides or binding fragment thereof in any of a variety of drug screening techniques. The PRO polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the PRO polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between PRO polypeptide or a fragment and the agent being tested. Alternatively, one can examine the diminution in complex formation between the PRO polypeptide and its target cell or target receptors caused by the agent being tested.

Thus, the present invention provides methods of screening for drugs or any other agents which can affect a PRO polypeptide-associated disease or disorder. These methods comprise contacting such an agent with an PRO polypeptide or fragment thereof and assaying (i) for the presence of a complex between the agent and the PRO polypeptide or fragment, or (ii) for the presence of a complex between the PRO polypeptide or fragment

and the cell, by methods well known in the art. In such competitive binding assays, the PRO polypeptide or fragment is typically labeled. After suitable incubation, free PRO polypeptide or fragment is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular agent to bind to PRO polypeptide or to interfere with the PRO polypeptide/cell complex.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to a polypeptide and is described in detail in WO 84/03564, published on September 13, 1984. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. As applied to a PRO polypeptide, the peptide test compounds are reacted with PRO polypeptide and washed. Bound PRO polypeptide is detected by methods well known in the art. Purified PRO polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the peptide and immobilize it on the solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding PRO polypeptide specifically compete with a test compound for binding to PRO polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PRO polypeptide.

EXAMPLE 13: Rational Drug Design

The goal of rational drug design is to produce structural analogs of biologically active polypeptide of interest (*i.e.*, a PRO polypeptide) or of small molecules with which they interact, *e.g.*, agonists, antagonists, or inhibitors. Any of these examples can be used to fashion drugs which are more active or stable forms of the PRO polypeptide or which enhance or interfere with the function of the PRO polypeptide *in vivo* (*c.f.*, Hodgson, Bio/Technology, 9: 19-21 (1991)).

In one approach, the three-dimensional structure of the PRO polypeptide, or of an PRO polypeptide-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the PRO polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of the PRO polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous PRO polypeptide-like molecules or to identify efficient inhibitors. Useful examples of rational drug design may include molecules which have improved activity or stability as shown by Braxton and Wells, Biochemistry, 31:7796-7801 (1992) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda *et al.*, J. Biochem., 113:742-746 (1993).

It is also possible to isolate a target-specific antibody, selected by functional assay, as described above, and then to solve its crystal structure. This approach, in principle, yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then

be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated peptides would then act as the pharmacore.

By virtue of the present invention, sufficient amounts of the PRO polypeptide may be made available to perform such analytical studies as X-ray crystallography. In addition, knowledge of the PRO polypeptide amino acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

EXAMPLE 14: Identification of PRO Polypeptides That Stimulate TNF- α Release In Human Blood (Assay 128)

This assay shows that certain PRO polypeptides of the present invention act to stimulate the release of TNF- α in human blood. PRO polypeptides testing positive in this assay are useful for, among other things, research purposes where stimulation of the release of TNF- α would be desired and for the therapeutic treatment of conditions wherein enhanced TNF- α release would be beneficial. Specifically, 200 μ l of human blood supplemented with 50mM Hepes buffer (pH 7.2) is aliquoted per well in a 96 well test plate. To each well is then added 300 μ l of either the test PRO polypeptide in 50 mM Hepes buffer (at various concentrations) or 50 mM Hepes buffer alone (negative control) and the plates are incubated at 37°C for 6 hours. The samples are then centrifuged and 50 μ l of plasma is collected from each well and tested for the presence of TNF- α by ELISA assay. A positive in the assay is a higher amount of TNF- α in the PRO polypeptide treated samples as compared to the negative control samples.

The following PRO polypeptides tested positive in this assay: PRO195, PRO202, PRO215, PRO221, PRO217, PRO222, PRO198, PRO245, PRO172, PRO265, PRO266, PRO344, PRO337, PRO322, PRO1286, PRO1279, PRO1338 and PRO1343.

EXAMPLE 15: Detection of Polypeptides That Affect Glucose or FFA Uptake in Skeletal Muscle (Assay 106)

This assay is designed to determine whether PRO polypeptides show the ability to affect glucose or FFA uptake by skeletal muscle cells. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of disorders where either the stimulation or inhibition of glucose uptake by skeletal muscle would be beneficial including, for example, diabetes or hyper- or hypo-insulinemia.

In a 96 well format, PRO polypeptides to be assayed are added to primary rat differentiated skeletal muscle, and allowed to incubate overnight. Then fresh media with the PRO polypeptide and +/- insulin are added to the wells. The sample media is then monitored to determine glucose and FFA uptake by the skeletal muscle cells. The insulin will stimulate glucose and FFA uptake by the skeletal muscle, and insulin in media without the PRO polypeptide is used as a positive control, and a limit for scoring. As the PRO polypeptide being tested may either stimulate or inhibit glucose and FFA uptake, results are scored as positive in the assay if greater than 1.5 times or less than 0.5 times the insulin control.

The following PRO polypeptides tested positive as being capable of affecting glucose and/or FFA uptake by skeletal muscle in this assay: PRO182, PRO366, PRO198, PRO172 and PRO719.

EXAMPLE 16: Chondrocyte Re-differentiation Assay (Assay 110)

This assay shows that certain polypeptides of the invention act to induce redifferentiation of chondrocytes, therefore, are expected to be useful for the treatment of various bone and/or cartilage disorders such as, for example, sports injuries and arthritis. The assay is performed as follows. Porcine chondrocytes are isolated by overnight collagenase digestion of articular cartilage of metacarpophalangeal joints of 4-6 month old female pigs. The isolated cells are then seeded at 25,000 cells/cm² in Ham F-12 containing 10% FBS and 4 µg/ml gentamycin. The culture media is changed every third day and the cells are then seeded in 96 well plates at 5,000 cells/well in 100µl of the same media without serum and 100 µl of the test PRO polypeptide, 5 nM staurosporin (positive control) or medium alone (negative control) is added to give a final volume of 200 µl/well. After 5 days of incubation at 37°C, a picture of each well is taken and the differentiation state of the chondrocytes is determined. A positive result in the assay occurs when the redifferentiation of the chondrocytes is determined to be more similar to the positive control than the negative control.

The following polypeptide tested positive in this assay: PRO182, PRO366, PRO198 and PRO1868.

EXAMPLE 17: Chondrocyte Proliferation Assay (Assay 111)

This assay is designed to determine whether PRO polypeptides of the present invention show the ability to induce the proliferation and/or redifferentiation of chondrocytes in culture. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of various bone and/or cartilage disorders such as, for example, sports injuries and arthritis.

Porcine chondrocytes are isolated by overnight collagenase digestion of articular cartilage of the metacarpophalangeal joint of 4-6 month old female pigs. The isolated cells are then seeded at 25,000 cells/cm² in Ham F-12 containing 10% FBS and 4 µg/ml gentamycin. The culture media is changed every third day and the cells are reseeded to 25,000 cells/cm² every five days. On day 12, the cells are seeded in 96 well plates at 5,000 cells/well in 100µl of the same media without serum and 100 µl of either serum-free medium (negative control), staurosporin (final concentration of 5 nM; positive control) or the test PRO polypeptide are added to give a final volume of 200 µl/well. After 5 days at 37°C, 20 µl of Alamar blue is added to each well and the plates are incubated for an additional 3 hours at 37°C. The fluorescence is then measured in each well (Ex:530 nm; Em: 590 nm). The fluorescence of a plate containing 200 µl of the serum-free medium is measured to obtain the background. A positive result in the assay is obtained when the fluorescence of the PRO polypeptide treated sample is more like that of the positive control than the negative control.

The following PRO polypeptides tested positive in this assay: PRO202, PRO224, PRO172 and PRO1312.

EXAMPLE 18: Detection of PRO Polypeptides That Affect Glucose or FFA Uptake by Primary Rat Adipocytes (Assay 94)

This assay is designed to determine whether PRO polypeptides show the ability to affect glucose or FFA uptake by adipocyte cells. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of disorders where either the stimulation or inhibition of glucose uptake by adipocytes

would be beneficial including, for example, obesity, diabetes or hyper- or hypo-insulinemia.

In a 96 well format, PRO polypeptides to be assayed are added to primary rat adipocytes, and allowed to incubate overnight. Samples are taken at 4 and 16 hours and assayed for glycerol, glucose and FFA uptake. After the 16 hour incubation, insulin is added to the media and allowed to incubate for 4 hours. At this time, a sample is taken and glycerol, glucose and FFA uptake is measured. Media containing insulin without the PRO polypeptide is used as a positive reference control. As the PRO polypeptide being tested may either stimulate or inhibit glucose and FFA uptake, results are scored as positive in the assay if greater than 1.5 times or less than 0.5 times the insulin control.

The following PRO polypeptides tested positive as being capable of affecting glucose and/or FFA uptake in this assay: PRO202, PRO211, PRO344 and PRO1338.

EXAMPLE 19: Gene Expression in Bovine Pericytes (Assay 105)

This assay is designed to identify PRO polypeptides which activate gene expression in pericytes. Such polypeptides would be expected to be useful as growth factors and/or for situations where the activation of gene expression is desired or beneficial. Bovine pericytes are plated on 60mm culture dishes in growth media for 1 week. On day 1, various PRO polypeptides are diluted (1%) and incubated with the pericytes for 1, 4 and 24 hr. timepoints. The cells are harvested and the RNA isolated using TRI-Reagent following the included instructions. The RNA is then quantified by reading the 260/280 OD using a spectrophotometer. The gene expression analysis is done by TaqMan reactions using Perkin Elmer reagents and specially designed bovine probes and primers. Expression of the following genes is analyzed: GAPDH, beta-integrin, connective tissue growth factor (CTGF), ICAM-1, monocyte chemoattractant protein-1 (MCP-1), osteopontin, transforming growth factor-beta (TGF-beta), TGF-beta receptor, tissue inhibitor of metalloproteinase (TIMP), tissue factor (TF), VEGF- α , thrombospondin, VEGF- β , angiopoietin-2, and collagenase. Replicates are then averaged and the SD determined. The gene expression levels are then normalized to GAPDH. These are then normalized to the expression levels obtained with a protein (PIN32) which does not significantly induce gene expression in bovine pericytes when compared to untreated controls. Any PRO polypeptide that gives a gene expression level 2-fold or higher over the PIN32 control is considered a positive hit.

The following PRO polypeptides tested positive in this assay: PRO366.

EXAMPLE 20: Identification of PRO Polypeptides That Activate Pericytes (Assay 125)

This assay shows that certain polypeptides of the invention act to activate proliferation of pericyte cells and, therefore, are useful not only as diagnostic markers for particular types of pericyte-associated tumors but also for giving rise to antagonists which would be expected to be useful for the therapeutic treatment of pericyte-associated tumors. Such PRO polypeptides also would be expected to be useful as growth factors and/or for situations where the induction of cell proliferation is desired or beneficial. Activation of pericyte proliferation also correlates with the induction of angiogenesis and, as such, PRO polypeptides capable of inducing pericyte proliferation would be expected to be useful for the treatment of conditions where induced angiogenesis would be beneficial including, for example, wound healing, and the like. Specifically, on day 1, pericytes are received

from VEC Technologies, and all but 5 ml media is removed from the flask. On day 2, the pericytes are trypsinized, washed, spun and plated on 96 well plates. On day 7, the media is removed and the pericytes are treated with 100 μ l of either the specific PRO polypeptide or control treatments (positive control = DME+5% +/- PDGF @ 500ng/ μ l; negative control=PIN32, a polypeptide determined to have no significant effect on pericyte proliferation). C-fos and GAPDH gene expression levels are then determined and the replicates are averaged and the SD is determined. The c-fos values are normalized to GAPDH and the results are expressed as fold increase over PIN32. Anything providing at least a 2-fold or higher response as compared to the negative control is considered positive for the assay.

The following polypeptides tested positive in this assay: PRO366.

EXAMPLE 21: Ability of PRO Polypeptides to Stimulate the Release of Proteoglycans from Cartilage (Assay 97)

The ability of various PRO polypeptides to stimulate the release of proteoglycans from cartilage tissue was tested as follows.

The metacarpophalangeal joint of 4-6 month old pigs was aseptically dissected, and articular cartilage was removed by free hand slicing being careful to avoid the underlying bone. The cartilage was minced and cultured in bulk for 24 hours in a humidified atmosphere of 95% air, 5% CO₂ in serum free (SF) media (DME/F12 1:1) with 0.1 % BSA and 100U/ml penicillin and 100 μ g/ml streptomycin. After washing three times, approximately 100 mg of articular cartilage was aliquoted into micronics tubes and incubated for an additional 24 hours in the above SF media. PRO polypeptides were then added at 1 % either alone or in combination with 18 ng/ml interleukin-1 α , a known stimulator of proteoglycan release from cartilage tissue. The supernatant was then harvested and assayed for the amount of proteoglycans using the 1,9-dimethyl-methylene blue (DMB) colorimetric assay (Farndale and Buttle, *Biochem. Biophys. Acta* 883:173-177 (1985)). A positive result in this assay indicates that the test polypeptide will find use, for example, in the treatment of sports-related joint problems, articular cartilage defects, osteoarthritis or rheumatoid arthritis.

When various PRO polypeptides were tested in the above assay, the polypeptides demonstrated a marked ability to stimulate release of proteoglycans from cartilage tissue both basally and after stimulation with interleukin-1 α and at 24 and 72 hours after treatment, thereby indicating that these PRO polypeptides are useful for stimulating proteoglycan release from cartilage tissue. As such, these PRO polypeptides are useful for the treatment of sports-related joint problems, articular cartilage defects, osteoarthritis or rheumatoid arthritis. The polypeptides testing positive in this assay are : PRO216.

EXAMPLE 22: Proliferation of Rat Utricular Supporting Cells (Assay 54)

This assay shows that certain polypeptides of the invention act as potent mitogens for inner ear supporting cells which are auditory hair cell progenitors and, therefore, are useful for inducing the regeneration of auditory hair cells and treating hearing loss in mammals. The assay is performed as follows. Rat UEC-4 utricular epithelial cells are aliquoted into 96 well plates with a density of 3000 cells/well in 200 μ l of serum-containing medium at 33°C. The cells are cultured overnight and are then switched to serum-free medium at

37°C. Various dilutions of PRO polypeptides (or nothing for a control) are then added to the cultures and the cells are incubated for 24 hours. After the 24 hour incubation, ³H-thymidine (1 µCi/well) is added and the cells are then cultured for an additional 24 hours. The cultures are then washed to remove unincorporated radiolabel, the cells harvested and Cpm per well determined. Cpm of at least 30% or greater in the PRO polypeptide treated cultures as compared to the control cultures is considered a positive in the assay.

5 The following polypeptides tested positive in this assay: PRO172.

EXAMPLE 23: Stimulatory Activity in Mixed Lymphocyte Reaction (MLR) Assay (Assay 24)

10 This example shows that certain polypeptides of the invention are active as a stimulator of the proliferation of stimulated T-lymphocytes. Compounds which stimulate proliferation of lymphocytes are useful therapeutically where enhancement of an immune response is beneficial. A therapeutic agent may take the form of antagonists of the polypeptide of the invention, for example, murine-human chimeric, humanized or human antibodies against the polypeptide.

15 The basic protocol for this assay is described in Current Protocols in Immunology, unit 3.12; edited by J E Coligan, A M Kruisbeek, D H Marglies, E M Shevach, W Strober, National Institutes of Health, Published by John Wiley & Sons, Inc.

20 More specifically, in one assay variant, peripheral blood mononuclear cells (PBMC) are isolated from mammalian individuals, for example a human volunteer, by leukopheresis (one donor will supply stimulator PBMCs, the other donor will supply responder PBMCs). If desired, the cells are frozen in fetal bovine serum and DMSO after isolation. Frozen cells may be thawed overnight in assay media (37°C, 5% CO₂) and then washed and resuspended to 3x10⁶ cells/ml of assay media (RPMI; 10% fetal bovine serum, 1% penicillin/streptomycin, 1% glutamine, 1% HEPES, 1% non-essential amino acids, 1% pyruvate). The stimulator PBMCs are prepared by irradiating the cells (about 3000 Rads).

 The assay is prepared by plating in triplicate wells a mixture of:

25 100:1 of test sample diluted to 1% or to 0.1%,

 50 :1 of irradiated stimulator cells, and

 50 :1 of responder PBMC cells.

 100 microliters of cell culture media or 100 microliter of CD4-IgG is used as the control. The wells are then incubated at 37°C, 5% CO₂ for 4 days. On day 5, each well is pulsed with tritiated thymidine (1.0 mCi/well; Amersham). After 6 hours the cells are washed 3 times and then the uptake of the label is evaluated.

30 In another variant of this assay, PBMCs are isolated from the spleens of Balb/c mice and C57B6 mice. The cells are teased from freshly harvested spleens in assay media (RPMI; 10% fetal bovine serum, 1% penicillin/streptomycin, 1% glutamine, 1% HEPES, 1% non-essential amino acids, 1% pyruvate) and the PBMCs are isolated by overlaying these cells over Lympholyte M (Organon Teknika), centrifuging at 2000 rpm for 20 minutes, collecting and washing the mononuclear cell layer in assay media and resuspending the cells to 1x10⁷ cells/ml of assay media. The assay is then conducted as described above.

35 Positive increases over control are considered positive with increases of greater than or equal to 180% being preferred. However, any value greater than control indicates a stimulatory effect for the test protein.

The following PRO polypeptides tested positive in this assay: PRO344.

EXAMPLE 24: Pericyte c-Fos Induction (Assay 93)

This assay shows that certain polypeptides of the invention act to induce the expression of c-fos in pericyte cells and, therefore, are useful not only as diagnostic markers for particular types of pericyte-associated tumors but also for giving rise to antagonists which would be expected to be useful for the therapeutic treatment of pericyte-associated tumors. Induction of c-fos expression in pericytes is also indicative of the induction of angiogenesis and, as such, PRO polypeptides capable of inducing the expression of c-fos would be expected to be useful for the treatment of conditions where induced angiogenesis would be beneficial including, for example, wound healing, and the like. Specifically, on day 1, pericytes are received from VEC Technologies and all but 5 ml of media is removed from flask. On day 2, the pericytes are trypsinized, washed, spun and then plated onto 96 well plates. On day 7, the media is removed and the pericytes are treated with 100 μ l of PRO polypeptide test samples and controls (positive control = DME+5% serum +/- PDGF at 500 ng/ml; negative control = protein 32). Replicates are averaged and SD/CV are determined. Fold increase over Protein 32 (buffer control) value indicated by chemiluminescence units (RLU) luminometer reading verses frequency is plotted on a histogram. Two-fold above Protein 32 value is considered positive for the assay. ASY Matrix: Growth media = low glucose DMEM = 20% FBS + 1X pen strep + 1X fungizone. Assay Media = low glucose DMEM +5% FBS.

The following polypeptides tested positive in this assay: PRO301, PRO619, PRO1066 and PRO1265.

EXAMPLE 25: Cytokine Release Assay (Assay 120)

This assay is designed to determine whether PRO polypeptides of the present invention are capable of inducing the release of cytokines from peripheral blood mononuclear cells (PBMCs). PRO polypeptides capable of inducing the release of cytokines from PBMCs are useful from the treatment of conditions which would benefit from enhanced cytokine release and will be readily evident to those of ordinary skill in the art. Specifically, 1×10^6 cells/ml of peripheral blood mononuclear cells (PBMC) are cultured with 1% of a PRO polypeptide for 3 days in complete RPMI media. The supernatant is then harvested and tested for increased concentrations of various cytokines by ELISA as compared to a human IgG treated control. A positive in the assay is a 10-fold or greater increase in cytokine concentration in the PRO polypeptide treated sample as compared to the human IgG treated control.

The following polypeptides tested positive in this assay: PRO526 and PRO1343.

EXAMPLE 26: Inhibition of A-Peptide Binding to Factor VIIA (Assay 118)

This assay is designed to identify PRO polypeptides which are capable of inhibiting the binding of A-peptide to factor VIIA, thereby affecting the blood coagulation cascade. PRO polypeptides testing positive in this assay are expected to be useful for the treatment of conditions where alteration of the blood coagulation cascade would be beneficial including, for example, stroke, heart attack and various coagulation disorders. These PRO polypeptides are also useful for the identification of agonist and antagonist molecules which would

also be useful for treatment of those conditions.

Specifically, 384 well plates are coated with soluble factor VIIA and are incubated overnight at 4°C. The wells are then decanted and are blocked by the addition of 0.5% BSA for 1 hour. The wells are then washed and 20µl of biotinylated A-peptide and either various concentration of the PRO polypeptide (test) or nothing (negative control) are added to each well. The plates are then incubated for 1 hour at room temperature. The wells are again washed and then 40µl of streptavidin-europium is added to each well. The plates are then incubated for 30 minutes at room temperature and then washed. 40µl of a fluorescence enhancement solution is then added to each well, the plates incubated for 5 minutes at room temperature and each well is then read on Wallac Victor reader under europium delayed fluorescence settings. Percent inhibition of binding of the A-peptide to the factor VIIA is then determined (as compared to the negative control), wherein a positive in the assay is a percent inhibition of 30% or greater.

The following PRO polypeptides tested positive in this assay: PRO182.

EXAMPLE 27: Inhibition of Adipocyte Differentiation Assay (Assay 66)

This assay is designed to identify PRO polypeptides which are capable of inhibiting insulin-induced differentiation of adipocytes. PRO polypeptides testing positive in this assay would be expected to be useful for the treatment of conditions associated with obesity, diabetes, etc.

Specifically, 3T3-L1 cells are seeded into the wells of 96 well plates at 6×10^4 cells/well and allowed to grow to confluency for 7 days. At day 7, the cells are treated with various concentrations of the PRO polypeptide (or nothing for the negative control) in the presence of 1µg/ml insulin, 0.25×10^{-6} M dexamethasone and 0.5mM IBMX. The samples are then incubated at 37°C in 7% CO₂ for 2 days. After the incubation, the media is removed by aspiration and the cells are washed with PBS and re-exposed to the PRO polypeptide (or nothing for the negative control) and 1µg/ml insulin. After 5 days, the media is removed and replaced with fresh PRO polypeptide (or nothing for the negative control) and insulin. After 5 days, the cells are lysed and the cell lysate is assayed using Sigma's Triglyceride [INT] kit (Sigma procedure #336). A positive in the assay is 20% greater inhibition of adipocyte differentiation in the PRO polypeptide treated samples as compared to the negative control.

The following PRO polypeptides tested positive in this assay: PRO185 and PRO198.

EXAMPLE 28: HUVEC Stimulation by PRO Polypeptides (Assay 131)

This assay is designed to identify PRO polypeptides which are capable of stimulating the proliferation of HUVEC cells. PRO polypeptides testing positive in this assay would be expected to be useful for inducing angiogenesis for the treatment of conditions where angiogenesis would be beneficial including, for example, wound healing, and the like. Antagonists of these PRO polypeptides would be expected to be useful for inhibiting angiogenesis for the treatment of, for example, tumors, and the like.

Specifically, COSTAR® flat bottom black plates are treated with fibronectin for 20 minutes and then washed twice with PBS. HUVEC cells are then plated at 2000 cells/well in an appropriate growth medium. The plates are then incubated overnight and then the PRO polypeptide (1% final concentration), nothing (negative

control) or IL1 β (3.3 ng/ml final concentration; positive control) is added. The plates are again incubated overnight, stained with ICAM1-Cy5 and read on FMAT. A positive in the assay is a 2-fold or greater increase in fluorescence as compared to the positive control.

The following PRO polypeptides tested positive in this assay: PRO222.

5 EXAMPLE 29: Promotion of Chondrocyte Redifferentiation (Assay 129)

This assay is designed to determine whether PRO polypeptides of the present invention show the ability to induce the proliferation and/or redifferentiation of chondrocytes in culture. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of various bone and/or cartilage disorders such as, for example, sports injuries and arthritis.

10 Porcine chondrocytes are isolated by overnight collagenase digestion of articular cartilage of the metacarpophalangeal joint of 4-6 month old female pigs. The isolated cells are then seeded at 25,000 cells/cm² in Ham F-12 containing 10% FBS and 4 μ g/ml gentamycin. The culture media is changed every third day. On day 12, the cells are seeded in 96 well plates at 5,000 cells/well in 100 μ l of the same media without serum and 100 μ l of either serum-free medium (negative control), staurosporin (final concentration of 5 nM; positive control) or the test PRO polypeptide are added to give a final volume of 200 μ l/well. After 5 days at 37°C, 22
15 μ l of media containing 100 μ g/ml Hoechst 33342 and 50 μ g/ml 5-CFDA is added to each well and incubated for an additional 10 minutes at 37°C. A picture of the green fluorescence is taken for each well and the differentiation state of the chondrocytes is calculated by morphometric analysis. A positive result in the assay is obtained when the > 50% of the PRO polypeptide treated cells are differentiated (compared to the background
20 obtained by the negative control).

The following PRO polypeptides tested positive in this assay: PRO301.

EXAMPLE 30: Microarray Analysis to Detect Overexpression of PRO Polypeptides in Cancerous Tumors

25 Nucleic acid microarrays, often containing thousands of gene sequences, are useful for identifying differentially expressed genes in diseased tissues as compared to their normal counterparts. Using nucleic acid microarrays, test and control mRNA samples from test and control tissue samples are reverse transcribed and labeled to generate cDNA probes. The cDNA probes are then hybridized to an array of nucleic acids immobilized on a solid support. The array is configured such that the sequence and position of each member of the array is known. For example, a selection of genes known to be expressed in certain disease states may be
30 arrayed on a solid support. Hybridization of a labeled probe with a particular array member indicates that the sample from which the probe was derived expresses that gene. If the hybridization signal of a probe from a test (disease tissue) sample is greater than hybridization signal of a probe from a control (normal tissue) sample, the gene or genes overexpressed in the disease tissue are identified. The implication of this result is that an overexpressed protein in a diseased tissue is useful not only as a diagnostic marker for the presence of the disease
35 condition, but also as a therapeutic target for treatment of the disease condition.

The methodology of hybridization of nucleic acids and microarray technology is well known in the art. In the present example, the specific preparation of nucleic acids for hybridization and probes, slides, and

hybridization conditions are all detailed in U.S. Provisional Patent Application Serial No. 60/193,767, filed on March 31, 2000 and which is herein incorporated by reference.

In the present example, cancerous tumors derived from various human tissues were studied for PRO polypeptide-encoding gene expression relative to non-cancerous human tissue in an attempt to identify those PRO polypeptides which are overexpressed in cancerous tumors. Two sets of experimental data were generated. In one set, cancerous human colon tumor tissue and matched non-cancerous human colon tumor tissue from the same patient ("matched colon control") were obtained and analyzed for PRO polypeptide expression using the above described microarray technology. In the second set of data, cancerous human tumor tissue from any of a variety of different human tumors was obtained and compared to a "universal" epithelial control sample which was prepared by pooling non-cancerous human tissues of epithelial origin, including liver, kidney, and lung. mRNA isolated from the pooled tissues represents a mixture of expressed gene products from these different tissues. Microarray hybridization experiments using the pooled control samples generated a linear plot in a 2-color analysis. The slope of the line generated in a 2-color analysis was then used to normalize the ratios of (test:control detection) within each experiment. The normalized ratios from various experiments were then compared and used to identify clustering of gene expression. Thus, the pooled "universal control" sample not only allowed effective relative gene expression determinations in a simple 2-sample comparison, it also allowed multi-sample comparisons across several experiments.

In the present experiments, nucleic acid probes derived from the herein described PRO polypeptide-encoding nucleic acid sequences were used in the creation of the microarray and RNA from the tumor tissues listed above were used for the hybridization thereto. A value based upon the normalized ratio:experimental ratio was designated as a "cutoff ratio". Only values that were above this cutoff ratio were determined to be significant. Table 8 below shows the results of these experiments, demonstrating that various PRO polypeptides of the present invention are significantly overexpressed in various human tumor tissues as compared to a non-cancerous human tissue control. As described above, these data demonstrate that the PRO polypeptides of the present invention are useful not only as diagnostic markers for the presence of one or more cancerous tumors, but also serve as therapeutic targets for the treatment of those tumors.

Table 8

<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
PRO177	breast tumor	universal normal control
PRO177	liver tumor	universal normal control
PRO177	lung tumor	universal normal control
PRO3574	breast tumor	universal normal control
PRO3574	colon tumor	matched normal colon control
PRO1280	breast tumor	universal normal control
PRO1280	lung tumor	universal normal control
PRO4984	lung tumor	universal normal control
PRO4988	colon tumor	universal normal control
PRO4988	lung tumor	universal normal control
PRO305	lung tumor	universal normal control
PRO305	colon tumor	universal normal control
PRO1866	prostate tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO1866	lung tumor	universal normal control
	PRO1866	colon tumor	universal normal control
	PRO4996	breast tumor	universal normal control
5	PRO4996	lung tumor	universal normal control
	PRO4406	lung tumor	universal normal control
	PRO4406	colon tumor	universal normal control
	PRO1120	colon tumor	universal normal control
	PRO1120	breast tumor	universal normal control
10	PRO1120	rectal tumor	universal normal control
	PRO4990	lung tumor	universal normal control
	PRO738	cervical tumor	universal normal control
	PRO738	lung tumor	universal normal control
	PRO738	breast tumor	universal normal control
15	PRO3577	lung tumor	universal normal control
	PRO1879	breast tumor	universal normal control
	PRO1879	lung tumor	universal normal control
	PRO1879	colon tumor	universal normal control
	PRO1471	lung tumor	universal normal control
20	PRO1076	prostate tumor	universal normal control
	PRO1483	lung tumor	universal normal control
	PRO4985	rectal tumor	universal normal control
	PRO4985	colon tumor	universal normal control
	PRO4985	breast tumor	universal normal control
25	PRO4985	lung tumor	universal normal control
	PRO5000	lung tumor	universal normal control
	PRO1881	liver tumor	universal normal control
	PRO1881	lung tumor	universal normal control
	PRO1881	breast tumor	universal normal control
30	PRO4314	lung tumor	universal normal control
	PRO4314	breast tumor	universal normal control
	PRO4987	lung tumor	universal normal control
	PRO4313	lung tumor	universal normal control
	PRO4313	breast tumor	universal normal control
35	PRO4799	colon tumor	universal normal control
	PRO4995	liver tumor	universal normal control
	PRO4995	colon tumor	universal normal control
	PRO4995	colon tumor	matched normal colon control
	PRO1341	prostate tumor	universal normal control
40	PRO1341	lung tumor	universal normal control
	PRO1341	colon tumor	universal normal control
	PRO1341	colon tumor	matched normal colon control
	PRO1777	lung tumor	universal normal control
	PRO1777	colon tumor	matched normal colon control
45	PRO3580	lung tumor	universal normal control
	PRO3580	prostate tumor	universal normal control
	PRO1779	lung tumor	universal normal control
	PRO1779	colon tumor	universal normal control
	PRO1779	cervical tumor	universal normal control
50	PRO1754	breast tumor	universal normal control
	PRO1754	lung tumor	universal normal control
	PRO1906	breast tumor	universal normal control
	PRO1906	colon tumor	universal normal control
	PRO1906	prostate tumor	universal normal control
55	PRO1870	breast tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO4329	lung tumor	universal normal control
	PRO4979	colon tumor	universal normal control
	PRO1885	rectal tumor	universal normal control
5	PRO1885	colon tumor	universal normal control
	PRO1885	colon tumor	matched normal colon control
	PRO1882	prostate tumor	universal normal control
	PRO1882	lung tumor	universal normal control
	PRO1882	colon tumor	universal normal control
10	PRO1882	breast tumor	universal normal control
	PRO1882	cervical tumor	universal normal control
	PRO4989	rectal tumor	universal normal control
	PRO4989	breast tumor	universal normal control
	PRO4989	colon tumor	matched normal colon control
15	PRO4989	colon tumor	universal normal control
	PRO4323	lung tumor	universal normal control
	PRO4323	liver tumor	universal normal control
	PRO1886	breast tumor	universal normal control
	PRO1886	lung tumor	universal normal control
20	PRO1886	rectal tumor	universal normal control
	PRO4395	colon tumor	universal normal control
	PRO4395	prostate tumor	universal normal control
	PRO4395	lung tumor	universal normal control
	PRO4395	cervical tumor	universal normal control
25	PRO1782	colon tumor	universal normal control
	PRO1782	lung tumor	universal normal control
	PRO4388	lung tumor	universal normal control
	PRO4341	breast tumor	universal normal control
	PRO4341	lung tumor	universal normal control
30	PRO3438	lung tumor	universal normal control
	PRO4321	breast tumor	universal normal control
	PRO4321	lung tumor	universal normal control
	PRO4321	colon tumor	universal normal control
	PRO4304	breast tumor	universal normal control
35	PRO4304	lung tumor	universal normal control
	PRO4403	colon tumor	universal normal control
	PRO4403	breast tumor	universal normal control
	PRO4403	lung tumor	universal normal control
	PRO4324	lung tumor	universal normal control
40	PRO4324	breast tumor	universal normal control
	PRO4303	cervical tumor	universal normal control
	PRO4303	lung tumor	universal normal control
	PRO4303	breast tumor	universal normal control
	PRO4303	colon tumor	universal normal control
45	PRO4303	prostate tumor	universal normal control
	PRO4305	breast tumor	universal normal control
	PRO4305	lung tumor	universal normal control
	PRO4305	colon tumor	universal normal control
	PRO4305	liver tumor	universal normal control
50	PRO4404	lung tumor	universal normal control
	PRO4404	breast tumor	universal normal control
	PRO4404	rectal tumor	universal normal control
	PRO1884	lung tumor	universal normal control
	PRO4349	colon tumor	universal normal control
55	PRO4349	lung tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO4401	colon tumor	universal normal control
	PRO4401	lung tumor	universal normal control
	PRO1867	lung tumor	universal normal control
5	PRO1867	liver tumor	universal normal control
	PRO4319	breast tumor	universal normal control
	PRO4319	lung tumor	universal normal control
	PRO4991	lung tumor	universal normal control
	PRO4991	colon tumor	universal normal control
10	PRO4398	lung tumor	universal normal control
	PRO4346	lung tumor	universal normal control
	PRO4350	colon tumor	universal normal control
	PRO4350	prostate tumor	universal normal control
	PRO4350	lung tumor	universal normal control
15	PRO4318	prostate tumor	universal normal control
	PRO4318	lung tumor	universal normal control
	PRO4340	breast tumor	universal normal control
	PRO4340	lung tumor	universal normal control
	PRO4400	breast tumor	universal normal control
20	PRO4400	lung tumor	universal normal control
	PRO4320	lung tumor	universal normal control
	PRO4409	lung tumor	universal normal control
	PRO4409	cervical tumor	universal normal control
	PRO4409	colon tumor	universal normal control
25	PRO4399	lung tumor	universal normal control
	PRO4399	breast tumor	universal normal control
	PRO4418	lung tumor	universal normal control
	PRO4418	breast tumor	universal normal control
	PRO4330	cervical tumor	universal normal control
30	PRO4330	colon tumor	matched normal colon control
	PRO4339	breast tumor	universal normal control
	PRO4339	colon tumor	universal normal control
	PRO4326	lung tumor	universal normal control
	PRO4326	colon tumor	universal normal control
35	PRO6014	breast tumor	universal normal control
	PRO3446	colon tumor	universal normal control
	PRO3446	lung tumor	universal normal control
	PRO4322	lung tumor	universal normal control
	PRO4322	rectal tumor	universal normal control
40	PRO4322	colon tumor	matched normal colon control
	PRO4381	breast tumor	universal normal control
	PRO4381	lung tumor	universal normal control
	PRO4381	colon tumor	universal normal control
	PRO4348	lung tumor	universal normal control
45	PRO4348	prostate tumor	universal normal control
	PRO4371	breast tumor	universal normal control
	PRO3742	colon tumor	universal normal control
	PRO3742	lung tumor	universal normal control
	PRO5773	lung tumor	universal normal control
50	PRO5773	colon tumor	universal normal control
	PRO5773	prostate tumor	universal normal control
	PRO5774	colon tumor	universal normal control
	PRO4343	colon tumor	universal normal control
	PRO4325	lung tumor	universal normal control
55	PRO4347	lung tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO4347	colon tumor	universal normal control
	PRO4347	rectal tumor	universal normal control
	PRO3743	colon tumor	universal normal control
5	PRO3743	lung tumor	universal normal control
	PRO3743	prostate tumor	universal normal control
	PRO4426	colon tumor	universal normal control
	PRO4500	colon tumor	universal normal control
	PRO4389	breast tumor	universal normal control
10	PRO4389	lung tumor	universal normal control
	PRO4337	colon tumor	universal normal control
	PRO4337	breast tumor	universal normal control
	PRO4337	lung tumor	universal normal control
	PRO4992	lung tumor	universal normal control
15	PRO5996	lung tumor	universal normal control
	PRO4345	lung tumor	universal normal control
	PRO4345	colon tumor	universal normal control
	PRO5780	lung tumor	universal normal control
	PRO5780	breast tumor	universal normal control
20	PRO5992	lung tumor	universal normal control
	PRO5992	colon tumor	universal normal control
	PRO5992	breast tumor	universal normal control
	PRO4428	prostate tumor	universal normal control
	PRO4994	lung tumor	universal normal control
25	PRO5995	lung tumor	universal normal control
	PRO5995	colon tumor	universal normal control
	PRO6094	lung tumor	universal normal control
	PRO6094	colon tumor	universal normal control
	PRO4317	lung tumor	universal normal control
30	PRO4317	colon tumor	universal normal control
	PRO4317	liver tumor	universal normal control
	PRO4317	colon tumor	matched normal colon control
	PRO5997	colon tumor	universal normal control
	PRO5997	lung tumor	universal normal control
35	PRO5005	lung tumor	universal normal control
	PRO5005	colon tumor	universal normal control
	PRO5004	colon tumor	universal normal control
	PRO6001	breast tumor	universal normal control
	PRO6013	colon tumor	universal normal control
40	PRO4502	lung tumor	universal normal control
	PRO4502	colon tumor	universal normal control
	PRO6007	breast tumor	universal normal control
	PRO6028	breast tumor	universal normal control
	PRO6028	colon tumor	universal normal control
45	PRO4327	prostate tumor	universal normal control
	PRO4315	colon tumor	universal normal control
	PRO5993	lung tumor	universal normal control
	PRO5993	colon tumor	universal normal control
	PRO4503	colon tumor	universal normal control
50	PRO4976	lung tumor	universal normal control
	PRO5798	lung tumor	universal normal control
	PRO5798	colon tumor	universal normal control
	PRO6242	colon tumor	universal normal control
	PRO6242	colon tumor	matched normal colon control
55	PRO6242	breast tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO6242	liver tumor	universal normal control
	PRO6242	rectal tumor	universal normal control
	PRO6095	breast tumor	universal normal control
5	PRO6095	lung tumor	universal normal control
	PRO6093	colon tumor	universal normal control
	PRO6093	breast tumor	universal normal control
	PRO6093	lung tumor	universal normal control
	PRO6093	colon tumor	matched normal colon control
10	PRO6012	colon tumor	universal normal control
	PRO6027	lung tumor	universal normal control
	PRO6027	colon tumor	universal normal control
	PRO6027	rectal tumor	universal normal control
	PRO6181	prostate tumor	universal normal control
15	PRO6181	lung tumor	universal normal control
	PRO6181	colon tumor	universal normal control
	PRO6097	colon tumor	universal normal control
	PRO6097	lung tumor	universal normal control
	PRO6090	lung tumor	universal normal control
20	PRO7171	lung tumor	universal normal control
	PRO7171	colon tumor	universal normal control
	PRO7171	breast tumor	universal normal control
	PRO6258	prostate tumor	universal normal control
	PRO6258	breast tumor	universal normal control
25	PRO6258	cervical tumor	universal normal control
	PRO6258	liver tumor	universal normal control
	PRO6258	colon tumor	universal normal control
	PRO9820	prostate tumor	universal normal control
	PRO6243	lung tumor	universal normal control
30	PRO6182	lung tumor	universal normal control
	PRO6079	lung tumor	universal normal control
	PRO6079	colon tumor	universal normal control
	PRO6079	breast tumor	universal normal control
	PRO6079	prostate tumor	universal normal control
35	PRO7434	lung tumor	universal normal control
	PRO9865	colon tumor	universal normal control
	PRO9828	colon tumor	universal normal control
	PRO196	colon tumor	universal normal control
	PRO196	lung tumor	universal normal control
40	PRO196	breast tumor	universal normal control
	PRO197	colon tumor	universal normal control
	PRO197	lung tumor	universal normal control
	PRO197	breast tumor	universal normal control
	PRO195	colon tumor	universal normal control
45	PRO195	lung tumor	universal normal control
	PRO195	breast tumor	universal normal control
	PRO187	lung tumor	universal normal control
	PRO187	liver tumor	universal normal control
	PRO182	colon tumor	universal normal control
50	PRO182	lung tumor	universal normal control
	PRO182	breast tumor	universal normal control
	PRO188	rectal tumor	universal normal control
	PRO183	colon tumor	universal normal control
	PRO183	lung tumor	universal normal control
55	PRO183	breast tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO183	rectal tumor	universal normal control
	PRO184	lung tumor	universal normal control
	PRO184	breast tumor	universal normal control
5	PRO185	lung tumor	universal normal control
	PRO200	colon tumor	universal normal control
	PRO200	lung tumor	universal normal control
	PRO200	breast tumor	universal normal control
	PRO200	rectal tumor	universal normal control
10	PRO202	colon tumor	universal normal control
	PRO202	lung tumor	universal normal control
	PRO202	breast tumor	universal normal control
	PRO202	rectal tumor	universal normal control
	PRO202	liver tumor	universal normal control
15	PRO214	colon tumor	universal normal control
	PRO214	lung tumor	universal normal control
	PRO215	colon tumor	universal normal control
	PRO215	lung tumor	universal normal control
	PRO215	breast tumor	universal normal control
20	PRO219	colon tumor	universal normal control
	PRO219	lung tumor	universal normal control
	PRO219	breast tumor	universal normal control
	PRO219	liver tumor	universal normal control
	PRO211	lung tumor	universal normal control
25	PRO211	breast tumor	universal normal control
	PRO220	colon tumor	universal normal control
	PRO220	lung tumor	universal normal control
	PRO220	breast tumor	universal normal control
	PRO366	colon tumor	universal normal control
30	PRO366	lung tumor	universal normal control
	PRO366	breast tumor	universal normal control
	PRO216	lung tumor	universal normal control
	PRO221	colon tumor	universal normal control
	PRO221	lung tumor	universal normal control
35	PRO221	breast tumor	universal normal control
	PRO228	lung tumor	universal normal control
	PRO228	breast tumor	universal normal control
	PRO217	lung tumor	universal normal control
	PRO217	breast tumor	universal normal control
40	PRO222	colon tumor	universal normal control
	PRO222	lung tumor	universal normal control
	PRO222	breast tumor	universal normal control
	PRO224	colon tumor	universal normal control
	PRO224	lung tumor	universal normal control
45	PRO224	breast tumor	universal normal control
	PRO224	prostate tumor	universal normal control
	PRO224	rectal tumor	universal normal control
	PRO230	colon tumor	universal normal control
	PRO230	lung tumor	universal normal control
50	PRO230	breast tumor	universal normal control
	PRO230	prostate tumor	universal normal control
	PRO198	colon tumor	universal normal control
	PRO198	lung tumor	universal normal control
	PRO198	breast tumor	universal normal control
55	PRO198	liver tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO226	lung tumor	universal normal control
	PRO226	breast tumor	universal normal control
	PRO261	lung tumor	universal normal control
5	PRO242	colon tumor	universal normal control
	PRO242	lung tumor	universal normal control
	PRO242	breast tumor	universal normal control
	PRO227	colon tumor	universal normal control
	PRO227	lung tumor	universal normal control
10	PRO237	colon tumor	universal normal control
	PRO237	lung tumor	universal normal control
	PRO237	breast tumor	universal normal control
	PRO237	prostate tumor	universal normal control
	PRO241	colon tumor	universal normal control
15	PRO241	lung tumor	universal normal control
	PRO241	breast tumor	universal normal control
	PRO231	colon tumor	universal normal control
	PRO231	lung tumor	universal normal control
	PRO231	breast tumor	universal normal control
20	PRO231	rectal tumor	universal normal control
	PRO235	colon tumor	universal normal control
	PRO235	lung tumor	universal normal control
	PRO235	breast tumor	universal normal control
	PRO235	liver tumor	universal normal control
25	PRO323	lung tumor	universal normal control
	PRO323	breast tumor	universal normal control
	PRO323	rectal tumor	universal normal control
	PRO245	colon tumor	universal normal control
	PRO245	lung tumor	universal normal control
30	PRO245	breast tumor	universal normal control
	PRO245	cervical tumor	universal normal control
	PRO245	liver tumor	universal normal control
	PRO246	colon tumor	universal normal control
	PRO246	lung tumor	universal normal control
35	PRO246	breast tumor	universal normal control
	PRO288	lung tumor	universal normal control
	PRO288	breast tumor	universal normal control
	PRO248	lung tumor	universal normal control
	PRO248	rectal tumor	universal normal control
40	PRO257	colon tumor	universal normal control
	PRO257	lung tumor	universal normal control
	PRO257	prostate tumor	universal normal control
	PRO172	colon tumor	universal normal control
	PRO172	lung tumor	universal normal control
45	PRO172	breast tumor	universal normal control
	PRO258	colon tumor	universal normal control
	PRO258	lung tumor	universal normal control
	PRO258	breast tumor	universal normal control
	PRO265	lung tumor	universal normal control
50	PRO265	breast tumor	universal normal control
	PRO265	rectal tumor	universal normal control
	PRO326	colon tumor	universal normal control
	PRO326	lung tumor	universal normal control
	PRO326	breast tumor	universal normal control
55	PRO326	liver tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO266	colon tumor	universal normal control
	PRO266	lung tumor	universal normal control
	PRO266	breast tumor	universal normal control
5	PRO269	lung tumor	universal normal control
	PRO269	rectal tumor	universal normal control
	PRO285	colon tumor	universal normal control
	PRO285	lung tumor	universal normal control
	PRO285	breast tumor	universal normal control
10	PRO328	colon tumor	universal normal control
	PRO328	lung tumor	universal normal control
	PRO328	breast tumor	universal normal control
	PRO344	breast tumor	universal normal control
	PRO272	lung tumor	universal normal control
15	PRO301	colon tumor	universal normal control
	PRO301	lung tumor	universal normal control
	PRO301	breast tumor	universal normal control
	PRO331	colon tumor	universal normal control
	PRO331	lung tumor	universal normal control
20	PRO331	breast tumor	universal normal control
	PRO332	colon tumor	universal normal control
	PRO332	lung tumor	universal normal control
	PRO332	breast tumor	universal normal control
	PRO353	colon tumor	universal normal control
25	PRO353	lung tumor	universal normal control
	PRO353	breast tumor	universal normal control
	PRO310	colon tumor	universal normal control
	PRO310	lung tumor	universal normal control
	PRO310	breast tumor	universal normal control
30	PRO310	rectal tumor	universal normal control
	PRO337	colon tumor	universal normal control
	PRO337	lung tumor	universal normal control
	PRO337	breast tumor	universal normal control
	PRO346	lung tumor	universal normal control
35	PRO350	lung tumor	universal normal control
	PRO350	breast tumor	universal normal control
	PRO526	colon tumor	universal normal control
	PRO526	lung tumor	universal normal control
	PRO526	breast tumor	universal normal control
40	PRO381	colon tumor	universal normal control
	PRO381	lung tumor	universal normal control
	PRO381	breast tumor	universal normal control
	PRO381	prostate tumor	universal normal control
	PRO846	colon tumor	universal normal control
45	PRO846	lung tumor	universal normal control
	PRO363	colon tumor	universal normal control
	PRO363	lung tumor	universal normal control
	PRO365	lung tumor	universal normal control
	PRO365	breast tumor	universal normal control
50	PRO1310	breast tumor	universal normal control
	PRO731	colon tumor	universal normal control
	PRO731	lung tumor	universal normal control
	PRO731	breast tumor	universal normal control
	PRO322	colon tumor	universal normal control
55	PRO322	lung tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO322	breast tumor	universal normal control
	PRO322	rectal tumor	universal normal control
	PRO322	liver tumor	universal normal control
5	PRO536	lung tumor	universal normal control
	PRO536	breast tumor	universal normal control
	PRO536	liver tumor	universal normal control
	PRO719	colon tumor	universal normal control
	PRO719	lung tumor	universal normal control
10	PRO719	breast tumor	universal normal control
	PRO619	colon tumor	universal normal control
	PRO619	lung tumor	universal normal control
	PRO619	breast tumor	universal normal control
	PRO771	colon tumor	universal normal control
15	PRO771	lung tumor	universal normal control
	PRO771	breast tumor	universal normal control
	PRO1083	colon tumor	universal normal control
	PRO1083	lung tumor	universal normal control
	PRO1083	breast tumor	universal normal control
20	PRO1083	prostate tumor	universal normal control
	PRO862	colon tumor	universal normal control
	PRO862	lung tumor	universal normal control
	PRO862	breast tumor	universal normal control
	PRO733	colon tumor	universal normal control
25	PRO733	lung tumor	universal normal control
	PRO733	breast tumor	universal normal control
	PRO733	liver tumor	universal normal control
	PRO1188	lung tumor	universal normal control
	PRO1188	breast tumor	universal normal control
30	PRO1188	rectal tumor	universal normal control
	PRO770	lung tumor	universal normal control
	PRO770	breast tumor	universal normal control
	PRO1080	colon tumor	universal normal control
	PRO1080	lung tumor	universal normal control
35	PRO1080	breast tumor	universal normal control
	PRO1017	colon tumor	universal normal control
	PRO1017	lung tumor	universal normal control
	PRO1017	breast tumor	universal normal control
	PRO1016	colon tumor	universal normal control
40	PRO1016	lung tumor	universal normal control
	PRO1016	breast tumor	universal normal control
	PRO1016	rectal tumor	universal normal control
	PRO792	lung tumor	universal normal control
	PRO938	colon tumor	universal normal control
45	PRO938	lung tumor	universal normal control
	PRO938	breast tumor	universal normal control
	PRO1012	colon tumor	universal normal control
	PRO1012	lung tumor	universal normal control
	PRO1012	rectal tumor	universal normal control
50	PRO1012	liver tumor	universal normal control
	PRO1008	lung tumor	universal normal control
	PRO1075	colon tumor	universal normal control
	PRO1075	lung tumor	universal normal control
	PRO1007	colon tumor	universal normal control
55	PRO1007	lung tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO1007	breast tumor	universal normal control
	PRO1007	rectal tumor	universal normal control
	PRO1056	colon tumor	universal normal control
5	PRO1056	lung tumor	universal normal control
	PRO1056	breast tumor	universal normal control
	PRO791	colon tumor	universal normal control
	PRO791	lung tumor	universal normal control
	PRO791	breast tumor	universal normal control
10	PRO791	rectal tumor	universal normal control
	PRO1111	colon tumor	universal normal control
	PRO1111	lung tumor	universal normal control
	PRO1111	breast tumor	universal normal control
	PRO812	lung tumor	universal normal control
15	PRO812	breast tumor	universal normal control
	PRO812	rectal tumor	universal normal control
	PRO1066	lung tumor	universal normal control
	PRO1185	colon tumor	universal normal control
	PRO1185	lung tumor	universal normal control
20	PRO1185	breast tumor	universal normal control
	PRO1031	lung tumor	universal normal control
	PRO1360	lung tumor	universal normal control
	PRO1360	breast tumor	universal normal control
	PRO1309	lung tumor	universal normal control
25	PRO1309	breast tumor	universal normal control
	PRO1107	lung tumor	universal normal control
	PRO1107	breast tumor	universal normal control
	PRO836	colon tumor	universal normal control
	PRO836	lung tumor	universal normal control
30	PRO1132	lung tumor	universal normal control
	PRO1132	breast tumor	universal normal control
	PRO1131	colon tumor	universal normal control
	PRO1131	lung tumor	universal normal control
	PRO1131	breast tumor	universal normal control
35	PRO1131	liver tumor	universal normal control
	PRO1130	colon tumor	universal normal control
	PRO1130	lung tumor	universal normal control
	PRO1130	breast tumor	universal normal control
	PRO844	colon tumor	universal normal control
40	PRO844	lung tumor	universal normal control
	PRO844	breast tumor	universal normal control
	PRO844	rectal tumor	universal normal control
	PRO1154	colon tumor	universal normal control
	PRO1154	lung tumor	universal normal control
45	PRO1154	rectal tumor	universal normal control
	PRO1154	liver tumor	universal normal control
	PRO1181	lung tumor	universal normal control
	PRO1181	breast tumor	universal normal control
	PRO1126	colon tumor	universal normal control
50	PRO1126	lung tumor	universal normal control
	PRO1126	breast tumor	universal normal control
	PRO1126	adrenal tumor	universal normal control
	PRO1186	colon tumor	universal normal control
	PRO1186	lung tumor	universal normal control
55	PRO1186	breast tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO1186	liver tumor	universal normal control
	PRO1198	colon tumor	universal normal control
	PRO1198	lung tumor	universal normal control
5	PRO1159	lung tumor	universal normal control
	PRO1159	breast tumor	universal normal control
	PRO1159	liver tumor	universal normal control
	PRO1265	colon tumor	universal normal control
	PRO1265	breast tumor	universal normal control
10	PRO1250	colon tumor	universal normal control
	PRO1250	lung tumor	universal normal control
	PRO1250	breast tumor	universal normal control
	PRO1475	colon tumor	universal normal control
	PRO1475	breast tumor	universal normal control
15	PRO1312	colon tumor	universal normal control
	PRO1312	lung tumor	universal normal control
	PRO1312	breast tumor	universal normal control
	PRO1308	colon tumor	universal normal control
	PRO1308	lung tumor	universal normal control
20	PRO1308	liver tumor	universal normal control
	PRO1326	colon tumor	universal normal control
	PRO1325	lung tumor	universal normal control
	PRO1326	breast tumor	universal normal control
	PRO1192	colon tumor	universal normal control
25	PRO1192	lung tumor	universal normal control
	PRO1192	breast tumor	universal normal control
	PRO1246	colon tumor	universal normal control
	PRO1246	lung tumor	universal normal control
	PRO1246	breast tumor	universal normal control
30	PRO1246	prostate tumor	universal normal control
	PRO1356	colon tumor	universal normal control
	PRO1356	lung tumor	universal normal control
	PRO1356	breast tumor	universal normal control
	PRO1275	lung tumor	universal normal control
35	PRO1275	breast tumor	universal normal control
	PRO1274	lung tumor	universal normal control
	PRO1358	colon tumor	universal normal control
	PRO1358	lung tumor	universal normal control
	PRO1358	prostate tumor	universal normal control
40	PRO1286	colon tumor	universal normal control
	PRO1286	lung tumor	universal normal control
	PRO1286	prostate tumor	universal normal control
	PRO1286	rectal tumor	universal normal control
	PRO1294	colon tumor	universal normal control
45	PRO1294	lung tumor	universal normal control
	PRO1294	breast tumor	universal normal control
	PRO1294	rectal tumor	universal normal control
	PRO1273	lung tumor	universal normal control
	PRO1273	rectal tumor	universal normal control
50	PRO1279	colon tumor	universal normal control
	PRO1279	lung tumor	universal normal control
	PRO1195	lung tumor	universal normal control
	PRO1195	breast tumor	universal normal control
	PRO1271	lung tumor	universal normal control
55	PRO1271	breast tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO1271	liver tumor	universal normal control
	PRO1338	colon tumor	universal normal control
	PRO1338	lung tumor	universal normal control
5	PRO1338	breast tumor	universal normal control
	PRO1343	colon tumor	universal normal control
	PRO1343	lung tumor	universal normal control
	PRO1343	breast tumor	universal normal control
	PRO1343	rectal tumor	universal normal control
10	PRO1434	lung tumor	universal normal control
	PRO1418	lung tumor	universal normal control
	PRO1418	liver tumor	universal normal control
	PRO1387	colon tumor	universal normal control
	PRO1387	lung tumor	universal normal control
15	PRO1387	prostate tumor	universal normal control
	PRO1387	rectal tumor	universal normal control
	PRO1384	colon tumor	universal normal control
	PRO1384	lung tumor	universal normal control
20	PRO1565	colon tumor	universal normal control
	PRO1565	lung tumor	universal normal control
	PRO1565	prostate tumor	universal normal control
	PRO1474	colon tumor	universal normal control
	PRO1474	lung tumor	universal normal control
	PRO1474	breast tumor	universal normal control
25	PRO1474	rectal tumor	universal normal control
	PRO1917	colon tumor	universal normal control
	PRO1917	lung tumor	universal normal control
	PRO1917	breast tumor	universal normal control
30	PRO1787	colon tumor	universal normal control
	PRO1787	lung tumor	universal normal control
	PRO1787	breast tumor	universal normal control
	PRO1556	lung tumor	universal normal control
	PRO1556	breast tumor	universal normal control
35	PRO1561	colon tumor	universal normal control
	PRO1561	lung tumor	universal normal control
	PRO1561	rectal tumor	universal normal control
	PRO1693	colon tumor	universal normal control
	PRO1693	lung tumor	universal normal control
	PRO1693	breast tumor	universal normal control
40	PRO1868	lung tumor	universal normal control
	PRO1868	breast tumor	universal normal control
	PRO1890	colon tumor	universal normal control
	PRO1890	lung tumor	universal normal control
	PRO1890	breast tumor	universal normal control
45	PRO1890	prostate tumor	universal normal control
	PRO1887	colon tumor	universal normal control
	PRO1887	breast tumor	universal normal control
	PRO4353	lung tumor	universal normal control
	PRO4353	breast tumor	universal normal control
50	PRO1801	colon tumor	universal normal control
	PRO1801	lung tumor	universal normal control
	PRO4357	lung tumor	universal normal control
	PRO4357	breast tumor	universal normal control
55	PRO4302	colon tumor	universal normal control
	PRO4302	lung tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO4302	breast tumor	universal normal control
	PRO4302	prostate tumor	universal normal control
5	PRO5990	colon tumor	universal normal control
	PRO5990	lung tumor	universal normal control
	PRO5990	breast tumor	universal normal control

EXAMPLE 31: Identification of Receptor/Ligand Interactions

10 In this assay, various PRO polypeptides are tested for ability to bind to a panel of potential receptor or ligand molecules for the purpose of identifying receptor/ligand interactions. The identification of a ligand for a known receptor, a receptor for a known ligand or a novel receptor/ligand pair is useful for a variety of indications including, for example, targeting bioactive molecules (linked to the ligand or receptor) to a cell known to express the receptor or ligand, use of the receptor or ligand as a reagent to detect the presence of the ligand or receptor in a composition suspected of containing the same, wherein the composition may comprise

15 cells suspected of expressing the ligand or receptor, modulating the growth of or another biological or immunological activity of a cell known to express or respond to the receptor or ligand, modulating the immune response of cells or toward cells that express the receptor or ligand, allowing the preparation of agonists, antagonists and/or antibodies directed against the receptor or ligand which will modulate the growth of or a biological or immunological activity of a cell expressing the receptor or ligand, and various other indications

20 which will be readily apparent to the ordinarily skilled artisan.

The assay is performed as follows. A PRO polypeptide of the present invention suspected of being a ligand for a receptor is expressed as a fusion protein containing the Fc domain of human IgG (an immunoadhesin). Receptor-ligand binding is detected by allowing interaction of the immunoadhesin polypeptide with cells (e.g. Cos cells) expressing candidate PRO polypeptide receptors and visualization of bound

25 immunoadhesin with fluorescent reagents directed toward the Fc fusion domain and examination by microscope. Cells expressing candidate receptors are produced by transient transfection, in parallel, of defined subsets of a library of cDNA expression vectors encoding PRO polypeptides that may function as receptor molecules. Cells are then incubated for 1 hour in the presence of the PRO polypeptide immunoadhesin being tested for possible receptor binding. The cells are then washed and fixed with paraformaldehyde. The cells are then incubated with

30 fluorescent conjugated antibody directed against the Fc portion of the PRO polypeptide immunoadhesin (e.g. FITC conjugated goat anti-human-Fc antibody). The cells are then washed again and examined by microscope. A positive interaction is judged by the presence of fluorescent labeling of cells transfected with cDNA encoding a particular PRO polypeptide receptor or pool of receptors and an absence of similar fluorescent labeling of similarly prepared cells that have been transfected with other cDNA or pools of cDNA. If a defined pool of

35 cDNA expression vectors is judged to be positive for interaction with a PRO polypeptide immunoadhesin, the individual cDNA species that comprise the pool are tested individually (the pool is "broken down") to determine the specific cDNA that encodes a receptor able to interact with the PRO polypeptide immunoadhesin.

In another embodiment of this assay, an epitope-tagged potential ligand PRO polypeptide (e.g. 8 histidine "His" tag) is allowed to interact with a panel of potential receptor PRO polypeptide molecules that have

been expressed as fusions with the Fc domain of human IgG (immunoadhesins). Following a 1 hour co-incubation with the epitope tagged PRO polypeptide, the candidate receptors are each immunoprecipitated with protein A beads and the beads are washed. Potential ligand interaction is determined by western blot analysis of the immunoprecipitated complexes with antibody directed towards the epitope tag. An interaction is judged to occur if a band of the anticipated molecular weight of the epitope tagged protein is observed in the western blot analysis with a candidate receptor, but is not observed to occur with the other members of the panel of potential receptors.

Using these assays, the following receptor/ligand interactions have been herein identified:

- (1) PRO1801 binds to PRO1114 and PRO4978.
- (2) PRO100 binds to PRO1114.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

0-1	Form - PCT/RO/134 (EASY) Indications Relating to Deposited Microorganism(s) or Other Biological Material (PCT Rule 13bis)	
0-1-1	Prepared using	PCT-EASY Version 2.91 (updated 10.10.2000)
0-2	International Application No.	
0-3	Applicant's or agent's file reference	P3330R1
1	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
1-1	page	98
1-2	line	34
1-3	Identification of Deposit	
1-3-1	Name of depositary institution	American Type Culture Collection
1-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
1-3-3	Date of deposit	14 April 1998 (14.04.1998)
1-3-4	Accession Number	ATCC 209771
1-4	Additional Indications	NONE
1-5	Designated States for Which Indications are Made	all designated States
1-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
2	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
2-1	page	98
2-2	line	35
2-3	Identification of Deposit	
2-3-1	Name of depositary institution	American Type Culture Collection
2-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
2-3-3	Date of deposit	09 February 1999 (09.02.1999)
2-3-4	Accession Number	ATCC 203654
2-4	Additional Indications	NONE
2-5	Designated States for Which Indications are Made	all designated States
2-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
3	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
3-1	page	98
3-2	line	36

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

3-3	Identification of Deposit	
3-3-1	Name of depositary institution	American Type Culture Collection
3-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
3-3-3	Date of deposit	25 May 1999 (25.05.1999)
3-3-4	Accession Number	ATCC PTA-127
3-4	Additional Indications	NONE
3-5	Designated States for Which Indications are Made	all designated States
3-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
4	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
4-1	page	98
4-2	line	37
4-3	Identification of Deposit	
4-3-1	Name of depositary institution	American Type Culture Collection
4-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
4-3-3	Date of deposit	27 July 1999 (27.07.1999)
4-3-4	Accession Number	ATCC PTA-429
4-4	Additional Indications	NONE
4-5	Designated States for Which Indications are Made	all designated States
4-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
5	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
5-1	page	98
5-2	line	38
5-3	Identification of Deposit	
5-3-1	Name of depositary institution	American Type Culture Collection
5-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
5-3-3	Date of deposit	27 July 1999 (27.07.1999)
5-3-4	Accession Number	ATCC PTA-432
5-4	Additional Indications	NONE
5-5	Designated States for Which Indications are Made	all designated States
5-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

6	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
6-1	page	98
6-2	line	39
6-3	Identification of Deposit	
6-3-1	Name of depositary institution	American Type Culture Collection
6-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
6-3-3	Date of deposit	10 December 1997 (10.12.1997)
6-3-4	Accession Number	ATCC 209525
6-4	Additional Indications	
6-4		NONE
6-5	Designated States for Which Indications are Made	
6-5		all designated States
6-6	Separate Furnishing of Indications	
6-6	These indications will be submitted to the International Bureau later	NONE
7	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
7-1	page	99
7-2	line	2
7-3	Identification of Deposit	
7-3-1	Name of depositary institution	American Type Culture Collection
7-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
7-3-3	Date of deposit	12 January 1999 (12.01.1999)
7-3-4	Accession Number	ATCC 203577
7-4	Additional Indications	
7-4		NONE
7-5	Designated States for Which Indications are Made	
7-5		all designated States
7-6	Separate Furnishing of Indications	
7-6	These indications will be submitted to the International Bureau later	NONE
8	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
8-1	page	99
8-2	line	3
8-3	Identification of Deposit	
8-3-1	Name of depositary institution	American Type Culture Collection
8-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
8-3-3	Date of deposit	27 July 1999 (27.07.1999)
8-3-4	Accession Number	ATCC PTA-430
8-4	Additional Indications	
8-4		NONE
8-5	Designated States for Which Indications are Made	
8-5		all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

8-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
9	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
9-1	page	99
9-2	line	4
9-3	Identification of Deposit	
9-3-1	Name of depositary institution	American Type Culture Collection
9-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
9-3-3	Date of deposit	08 June 1999 (08.06.1999)
9-3-4	Accession Number	ATCC PTA-203
9-4	Additional Indications	NONE
9-5	Designated States for Which Indications are Made	all designated States
9-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
10	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
10-1	page	99
10-2	line	5
10-3	Identification of Deposit	
10-3-1	Name of depositary institution	American Type Culture Collection
10-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
10-3-3	Date of deposit	01 July 1998 (01.07.1998)
10-3-4	Accession Number	ATCC 203040
10-4	Additional Indications	NONE
10-5	Designated States for Which Indications are Made	all designated States
10-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
11	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
11-1	page	99
11-2	line	6
11-3	Identification of Deposit	
11-3-1	Name of depositary institution	American Type Culture Collection
11-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
11-3-3	Date of deposit	31 August 1999 (31.08.1999)
11-3-4	Accession Number	ATCC PTA-611

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

11-4	Additional Indications	NONE
11-5	Designated States for Which Indications are Made	all designated States
11-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
12	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
12-1	page	99
12-2	line	7
12-3	Identification of Deposit	
12-3-1	Name of depositary institution	American Type Culture Collection
12-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
12-3-3	Date of deposit	21 January 1998 (21.01.1998)
12-3-4	Accession Number	ATCC 209593
12-4	Additional Indications	NONE
12-5	Designated States for Which Indications are Made	all designated States
12-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
13	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
13-1	page	99
13-2	line	8
13-3	Identification of Deposit	
13-3-1	Name of depositary institution	American Type Culture Collection
13-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
13-3-3	Date of deposit	09 February 1999 (09.02.1999)
13-3-4	Accession Number	ATCC 203649
13-4	Additional Indications	NONE
13-5	Designated States for Which Indications are Made	all designated States
13-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
14	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
14-1	page	99
14-2	line	9

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

14-3	Identification of Deposit	
14-3-1	Name of depositary institution	American Type Culture Collection
14-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
14-3-3	Date of deposit	12 January 1999 (12.01.1999)
14-3-4	Accession Number	ATCC 203574
14-4	Additional Indications	NONE
14-5	Designated States for Which Indications are Made	all designated States
14-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
15	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
15-1	page	99
15-2	line	10
15-3	Identification of Deposit	
15-3-1	Name of depositary institution	American Type Culture Collection
15-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
15-3-3	Date of deposit	25 May 1999 (25.05.1999)
15-3-4	Accession Number	ATCC PTA-129
15-4	Additional Indications	NONE
15-5	Designated States for Which Indications are Made	all designated States
15-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
16	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
16-1	page	99
16-2	line	11
16-3	Identification of Deposit	
16-3-1	Name of depositary institution	American Type Culture Collection
16-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
16-3-3	Date of deposit	27 May 1998 (27.05.1998)
16-3-4	Accession Number	ATCC 209905
16-4	Additional Indications	NONE
16-5	Designated States for Which Indications are Made	all designated States
16-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

17	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
17-1	page	99
17-2	line	12
17-3	Identification of Deposit	
17-3-1	Name of depositary institution	American Type Culture Collection
17-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
17-3-3	Date of deposit	12 January 1999 (12.01.1999)
17-3-4	Accession Number	ATCC 203585
17-4	Additional Indications	NONE
17-5	Designated States for Which Indications are Made	all designated States
17-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
18	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
18-1	page	99
18-2	line	13
18-3	Identification of Deposit	
18-3-1	Name of depositary institution	American Type Culture Collection
18-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
18-3-3	Date of deposit	09 February 1999 (09.02.1999)
18-3-4	Accession Number	ATCC 203665
18-4	Additional Indications	NONE
18-5	Designated States for Which Indications are Made	all designated States
18-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
19	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
19-1	page	99
19-2	line	14
19-3	Identification of Deposit	
19-3-1	Name of depositary institution	American Type Culture Collection
19-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
19-3-3	Date of deposit	27 July 1999 (27.07.1999)
19-3-4	Accession Number	ATCC PTA-427
19-4	Additional Indications	NONE
19-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

19-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
20	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
20-1	page	99
20-2	line	15
20-3	Identification of Deposit	
20-3-1	Name of depositary institution	American Type Culture Collection
20-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
20-3-3	Date of deposit	31 August 1999 (31.08.1999)
20-3-4	Accession Number	ATCC PTA-615
20-4	Additional Indications	NONE
20-5	Designated States for Which Indications are Made	all designated States
20-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
21	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
21-1	page	99
21-2	line	16
21-3	Identification of Deposit	
21-3-1	Name of depositary institution	American Type Culture Collection
21-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
21-3-3	Date of deposit	12 January 1999 (12.01.1999)
21-3-4	Accession Number	ATCC 203582
21-4	Additional Indications	NONE
21-5	Designated States for Which Indications are Made	all designated States
21-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
22	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
22-1	page	99
22-2	line	17
22-3	Identification of Deposit	
22-3-1	Name of depositary institution	American Type Culture Collection
22-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
22-3-3	Date of deposit	09 March 1999 (09.03.1999)
22-3-4	Accession Number	ATCC 203838

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

22-4	Additional Indications	NONE
22-5	Designated States for Which Indications are Made	all designated States
22-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
23	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
23-1	page	99
23-2	line	18
23-3	Identification of Deposit	
23-3-1	Name of depositary institution	American Type Culture Collection
23-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
23-3-3	Date of deposit	27 July 1999 (27.07.1999)
23-3-4	Accession Number	ATCC PTA-428
23-4	Additional Indications	NONE
23-5	Designated States for Which Indications are Made	all designated States
23-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
24	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
24-1	page	99
24-2	line	19
24-3	Identification of Deposit	
24-3-1	Name of depositary institution	American Type Culture Collection
24-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
24-3-3	Date of deposit	09 March 1999 (09.03.1999)
24-3-4	Accession Number	ATCC 203836
24-4	Additional Indications	NONE
24-5	Designated States for Which Indications are Made	all designated States
24-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
25	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
25-1	page	99
25-2	line	20

PCT

P3330R1

Original (for SUBMISSION) - printed on 01:12:2000 02:57:35 PM

25-3	Identification of Deposit	
25-3-1	Name of depositary institution	American Type Culture Collection
25-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
25-3-3	Date of deposit	08 June 1999 (08.06.1999)
25-3-4	Accession Number	ATCC PTA-205
25-4	Additional Indications	NONE
25-5	Designated States for Which Indications are Made	all designated States
25-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
26	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
26-1	page	99
26-2	line	21
26-3	Identification of Deposit	
26-3-1	Name of depositary institution	American Type Culture Collection
26-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
26-3-3	Date of deposit	27 July 1999 (27.07.1999)
26-3-4	Accession Number	ATCC PTA-431
26-4	Additional Indications	NONE
26-5	Designated States for Which Indications are Made	all designated States
26-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
27	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
27-1	page	99
27-2	line	22
27-3	Identification of Deposit	
27-3-1	Name of depositary institution	American Type Culture Collection
27-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
27-3-3	Date of deposit	09 February 1999 (09.02.1999)
27-3-4	Accession Number	ATCC 203659
27-4	Additional Indications	NONE
27-5	Designated States for Which Indications are Made	all designated States
27-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

28	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
28-1	page	99
28-2	line	23
28-3	Identification of Deposit	
28-3-1	Name of depositary institution	American Type Culture Collection
28-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
28-3-3	Date of deposit	12 January 1999 (12.01.1999)
28-3-4	Accession Number	ATCC 203584
28-4	Additional Indications	NONE
28-5	Designated States for Which Indications are Made	all designated States
28-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
29	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
29-1	page	99
29-2	line	24
29-3	Identification of Deposit	
29-3-1	Name of depositary institution	American Type Culture Collection
29-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
29-3-3	Date of deposit	25 May 1999 (25.05.1999)
29-3-4	Accession Number	ATCC PTA-126
29-4	Additional Indications	NONE
29-5	Designated States for Which Indications are Made	all designated States
29-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
30	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
30-1	page	99
30-2	line	25
30-3	Identification of Deposit	
30-3-1	Name of depositary institution	American Type Culture Collection
30-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
30-3-3	Date of deposit	25 May 1999 (25.05.1999)
30-3-4	Accession Number	ATCC PTA-128
30-4	Additional Indications	NONE
30-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

30-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
31	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
31-1	page	99
31-2	line	26
31-3	Identification of Deposit	
31-3-1	Name of depositary institution	American Type Culture Collection
31-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
31-3-3	Date of deposit	09 February 1999 (09.02.1999)
31-3-4	Accession Number	ATCC 203664
31-4	Additional Indications	NONE
31-5	Designated States for Which Indications are Made	all designated States
31-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
32	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
32-1	page	99
32-2	line	27
32-3	Identification of Deposit	
32-3-1	Name of depositary institution	American Type Culture Collection
32-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
32-3-3	Date of deposit	12 January 1999 (12.01.1999)
32-3-4	Accession Number	ATCC 203578
32-4	Additional Indications	NONE
32-5	Designated States for Which Indications are Made	all designated States
32-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
33	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
33-1	page	99
33-2	line	28
33-3	Identification of Deposit	
33-3-1	Name of depositary institution	American Type Culture Collection
33-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
33-3-3	Date of deposit	22 December 1998 (22.12.1998)
33-3-4	Accession Number	ATCC 203554

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

33-4	Additional Indications	NONE
33-5	Designated States for Which Indications are Made	all designated States
33-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
34	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
34-1	page	99
34-2	line	29
34-3	Identification of Deposit	
34-3-1	Name of depositary institution	American Type Culture Collection
34-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
34-3-3	Date of deposit	16 March 1999 (16.03.1999)
34-3-4	Accession Number	ATCC 203850
34-4	Additional Indications	NONE
34-5	Designated States for Which Indications are Made	all designated States
34-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
35	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
35-1	page	99
35-2	line	30
35-3	Identification of Deposit	
35-3-1	Name of depositary institution	American Type Culture Collection
35-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
35-3-3	Date of deposit	11 May 1999 (11.05.1999)
35-3-4	Accession Number	ATCC PTA-45
35-4	Additional Indications	NONE
35-5	Designated States for Which Indications are Made	all designated States
35-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
36	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
36-1	page	99
36-2	line	31

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

36-3	Identification of Deposit	
36-3-1	Name of depositary institution	American Type Culture Collection
36-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
36-3-3	Date of deposit	22 December 1998 (22.12.1998)
36-3-4	Accession Number	ATCC 203545
36-4	Additional Indications	NONE
36-5	Designated States for Which Indications are Made	all designated States
36-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
37	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
37-1	page	99
37-2	line	32
37-3	Identification of Deposit	
37-3-1	Name of depositary institution	American Type Culture Collection
37-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
37-3-3	Date of deposit	22 December 1998 (22.12.1998)
37-3-4	Accession Number	ATCC 203544
37-4	Additional Indications	NONE
37-5	Designated States for Which Indications are Made	all designated States
37-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
38	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
38-1	page	99
38-2	line	33
38-3	Identification of Deposit	
38-3-1	Name of depositary institution	American Type Culture Collection
38-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
38-3-3	Date of deposit	15 June 1999 (15.06.1999)
38-3-4	Accession Number	ATCC PTA-234
38-4	Additional Indications	NONE
38-5	Designated States for Which Indications are Made	all designated States
38-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

39	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
39-1	page	99
39-2	line	34
39-3	Identification of Deposit	
39-3-1	Name of depositary institution	American Type Culture Collection
39-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
39-3-3	Date of deposit	16 March 1999 (16.03.1999)
39-3-4	Accession Number	ATCC 203848
39-4	Additional Indications	NONE
39-5	Designated States for Which Indications are Made	all designated States
39-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
40	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
40-1	page	99
40-2	line	35
40-3	Identification of Deposit	
40-3-1	Name of depositary institution	American Type Culture Collection
40-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
40-3-3	Date of deposit	22 December 1998 (22.12.1998)
40-3-4	Accession Number	ATCC 203555
40-4	Additional Indications	NONE
40-5	Designated States for Which Indications are Made	all designated States
40-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
41	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
41-1	page	99
41-2	line	36
41-3	Identification of Deposit	
41-3-1	Name of depositary institution	American Type Culture Collection
41-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
41-3-3	Date of deposit	20 April 1999 (20.04.1999)
41-3-4	Accession Number	ATCC 203949
41-4	Additional Indications	NONE
41-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

41-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
42	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
42-1	page	99
42-2	line	37
42-3	Identification of Deposit	
42-3-1	Name of depositary institution	American Type Culture Collection
42-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
42-3-3	Date of deposit	15 December 1998 (15.12.1998)
42-3-4	Accession Number	ATCC 203539
42-4	Additional Indications	NONE
42-5	Designated States for Which Indications are Made	all designated States
42-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
43	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
43-1	page	99
43-2	line	38
43-3	Identification of Deposit	
43-3-1	Name of depositary institution	American Type Culture Collection
43-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
43-3-3	Date of deposit	23 March 1999 (23.03.1999)
43-3-4	Accession Number	ATCC 203871
43-4	Additional Indications	NONE
43-5	Designated States for Which Indications are Made	all designated States
43-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
44	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
44-1	page	99
44-2	line	39
44-3	Identification of Deposit	
44-3-1	Name of depositary institution	American Type Culture Collection
44-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
44-3-3	Date of deposit	23 March 1999 (23.03.1999)
44-3-4	Accession Number	ATCC 203862

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

44-4	Additional Indications	NONE
44-5	Designated States for Which Indications are Made	all designated States
44-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
45	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
45-1	page	99
45-2	line	40
45-3	Identification of Deposit	
45-3-1	Name of depositary institution	American Type Culture Collection
45-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
45-3-3	Date of deposit	10 August 1999 (10.08.1999)
45-3-4	Accession Number	ATCC PTA-510
45-4	Additional Indications	NONE
45-5	Designated States for Which Indications are Made	all designated States
45-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
46	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
46-1	page	99
46-2	line	41
46-3	Identification of Deposit	
46-3-1	Name of depositary institution	American Type Culture Collection
46-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
46-3-3	Date of deposit	20 January 1999 (20.01.1999)
46-3-4	Accession Number	ATCC 203603
46-4	Additional Indications	NONE
46-5	Designated States for Which Indications are Made	all designated States
46-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
47	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
47-1	page	99
47-2	line	42

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

47-3	Identification of Deposit	
47-3-1	Name of depositary institution	American Type Culture Collection
47-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
47-3-3	Date of deposit	02 March 1999 (02.03.1999)
47-3-4	Accession Number	ATCC 203813
47-4	Additional Indications	NONE
47-5	Designated States for Which Indications are Made	all designated States
47-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
48	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
48-1	page	99
48-2	line	43
48-3	Identification of Deposit	
48-3-1	Name of depositary institution	American Type Culture Collection
48-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
48-3-3	Date of deposit	02 March 1999 (02.03.1999)
48-3-4	Accession Number	ATCC 203812
48-4	Additional Indications	NONE
48-5	Designated States for Which Indications are Made	all designated States
48-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
49	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
49-1	page	99
49-2	line	44
49-3	Identification of Deposit	
49-3-1	Name of depositary institution	American Type Culture Collection
49-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
49-3-3	Date of deposit	29 October 1998 (29.10.1998)
49-3-4	Accession Number	ATCC 203391
49-4	Additional Indications	NONE
49-5	Designated States for Which Indications are Made	all designated States
49-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

50	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
50-1	page	99
50-2	line	45
50-3	Identification of Deposit	
50-3-1	Name of depositary institution	American Type Culture Collection
50-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
50-3-3	Date of deposit	27 April 1999 (27.04.1999)
50-3-4	Accession Number	ATCC 203965
50-4	Additional Indications	NONE
50-5	Designated States for Which Indications are Made	all designated States
50-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
51	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
51-1	page	99
51-2	line	46
51-3	Identification of Deposit	
51-3-1	Name of depositary institution	American Type Culture Collection
51-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
51-3-3	Date of deposit	02 March 1999 (02.03.1999)
51-3-4	Accession Number	ATCC 203816
51-4	Additional Indications	NONE
51-5	Designated States for Which Indications are Made	all designated States
51-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
52	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
52-1	page	99
52-2	line	47
52-3	Identification of Deposit	
52-3-1	Name of depositary institution	American Type Culture Collection
52-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
52-3-3	Date of deposit	02 March 1999 (02.03.1999)
52-3-4	Accession Number	ATCC 203814
52-4	Additional Indications	NONE
52-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

52-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
53	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
53-1	page	99
53-2	line	48
53-3	Identification of Deposit	
53-3-1	Name of depositary institution	American Type Culture Collection
53-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
53-3-3	Date of deposit	02 March 1999 (02.03.1999)
53-3-4	Accession Number	ATCC 203810
53-4	Additional Indications	NONE
53-5	Designated States for Which Indications are Made	all designated States
53-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
54	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
54-1	page	99
54-2	line	49
54-3	Identification of Deposit	
54-3-1	Name of depositary institution	American Type Culture Collection
54-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
54-3-3	Date of deposit	04 May 1999 (04.05.1999)
54-3-4	Accession Number	ATCC PTA-22
54-4	Additional Indications	NONE
54-5	Designated States for Which Indications are Made	all designated States
54-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
55	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
55-1	page	99
55-2	line	50
55-3	Identification of Deposit	
55-3-1	Name of depositary institution	American Type Culture Collection
55-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
55-3-3	Date of deposit	12 January 1999 (12.01.1999)
55-3-4	Accession Number	ATCC 203580

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

55-4	Additional Indications	NONE
55-5	Designated States for Which Indications are Made	all designated States
55-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
56	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
56-1	page	99
56-2	line	51
56-3	Identification of Deposit	
56-3-1	Name of depositary institution	American Type Culture Collection
56-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
56-3-3	Date of deposit	30 March 1999 (30.03.1999)
56-3-4	Accession Number	ATCC 203889
56-4	Additional Indications	NONE
56-5	Designated States for Which Indications are Made	all designated States
56-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
57	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
57-1	page	99
57-2	line	52
57-3	Identification of Deposit	
57-3-1	Name of depositary institution	American Type Culture Collection
57-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
57-3-3	Date of deposit	27 April 1999 (27.04.1999)
57-3-4	Accession Number	ATCC 203964
57-4	Additional Indications	NONE
57-5	Designated States for Which Indications are Made	all designated States
57-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
58	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
58-1	page	99
58-2	line	53

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

58-3	Identification of Deposit	
58-3-1	Name of depositary institution	American Type Culture Collection
58-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
58-3-3	Date of deposit	22 December 1998 (22.12.1998)
58-3-4	Accession Number	ATCC 203548
58-4	Additional Indications	NONE
58-5	Designated States for Which Indications are Made	all designated States
58-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
59	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
59-1	page	99
59-2	line	54
59-3	Identification of Deposit	
59-3-1	Name of depositary institution	American Type Culture Collection
59-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
59-3-3	Date of deposit	02 March 1999 (02.03.1999)
59-3-4	Accession Number	ATCC 203817
59-4	Additional Indications	NONE
59-5	Designated States for Which Indications are Made	all designated States
59-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
60	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
60-1	page	99
60-2	line	55
60-3	Identification of Deposit	
60-3-1	Name of depositary institution	American Type Culture Collection
60-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
60-3-3	Date of deposit	15 June 1999 (15.06.1999)
60-3-4	Accession Number	ATCC PTA-235
60-4	Additional Indications	NONE
60-5	Designated States for Which Indications are Made	all designated States
60-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

61	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
61-1	page	100
61-2	line	2
61-3	Identification of Deposit	
61-3-1	Name of depositary institution	American Type Culture Collection
61-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
61-3-3	Date of deposit	27 April 1999 (27.04.1999)
61-3-4	Accession Number	ATCC 203968
61-4	Additional Indications	NONE
61-5	Designated States for Which Indications are Made	all designated States
61-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
62	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
62-1	page	100
62-2	line	3
62-3	Identification of Deposit	
62-3-1	Name of depositary institution	American Type Culture Collection
62-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
62-3-3	Date of deposit	30 March 1999 (30.03.1999)
62-3-4	Accession Number	ATCC 203894
62-4	Additional Indications	NONE
62-5	Designated States for Which Indications are Made	all designated States
62-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
63	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
63-1	page	100
63-2	line	4
63-3	Identification of Deposit	
63-3-1	Name of depositary institution	American Type Culture Collection
63-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
63-3-3	Date of deposit	30 March 1999 (30.03.1999)
63-3-4	Accession Number	ATCC 203893
63-4	Additional Indications	NONE
63-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

63-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
64	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
64-1	page	100
64-2	line	5
64-3	Identification of Deposit	
64-3-1	Name of depositary institution	American Type Culture Collection
64-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
64-3-3	Date of deposit	02 March 1999 (02.03.1999)
64-3-4	Accession Number	ATCC 203811
64-4	Additional Indications	NONE
64-5	Designated States for Which Indications are Made	all designated States
64-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
65	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
65-1	page	100
65-2	line	6
65-3	Identification of Deposit	
65-3-1	Name of depositary institution	American Type Culture Collection
65-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
65-3-3	Date of deposit	23 March 1999 (23.03.1999)
65-3-4	Accession Number	ATCC 203867
65-4	Additional Indications	NONE
65-5	Designated States for Which Indications are Made	all designated States
65-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
66	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
66-1	page	100
66-2	line	7
66-3	Identification of Deposit	
66-3-1	Name of depositary institution	American Type Culture Collection
66-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
66-3-3	Date of deposit	27 April 1999 (27.04.1999)
66-3-4	Accession Number	ATCC 203963

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

66-4	Additional Indications	NONE
66-5	Designated States for Which Indications are Made	all designated States
66-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
67	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
67-1	page	100
67-2	line	8
67-3	Identification of Deposit	
67-3-1	Name of depositary institution	American Type Culture Collection
67-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
67-3-3	Date of deposit	02 March 1999 (02.03.1999)
67-3-4	Accession Number	ATCC 203815
67-4	Additional Indications	NONE
67-5	Designated States for Which Indications are Made	all designated States
67-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
68	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
68-1	page	100
68-2	line	9
68-3	Identification of Deposit	
68-3-1	Name of depositary institution	American Type Culture Collection
68-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
68-3-3	Date of deposit	30 March 1999 (30.03.1999)
68-3-4	Accession Number	ATCC 203890
68-4	Additional Indications	NONE
68-5	Designated States for Which Indications are Made	all designated States
68-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
69	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
69-1	page	100
69-2	line	10

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

69-3	Identification of Deposit	
69-3-1	Name of depositary institution	American Type Culture Collection
69-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
69-3-3	Date of deposit	25 May 1999 (25.05.1999)
69-3-4	Accession Number	ATCC PTA-130
69-4	Additional Indications	NONE
69-5	Designated States for Which Indications are Made	all designated States
69-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
70	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
70-1	page	100
70-2	line	11
70-3	Identification of Deposit	
70-3-1	Name of depositary institution	American Type Culture Collection
70-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
70-3-3	Date of deposit	27 April 1999 (27.04.1999)
70-3-4	Accession Number	ATCC 203970
70-4	Additional Indications	NONE
70-5	Designated States for Which Indications are Made	all designated States
70-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
71	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
71-1	page	100
71-2	line	12
71-3	Identification of Deposit	
71-3-1	Name of depositary institution	American Type Culture Collection
71-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
71-3-3	Date of deposit	16 March 1999 (16.03.1999)
71-3-4	Accession Number	ATCC 203845
71-4	Additional Indications	NONE
71-5	Designated States for Which Indications are Made	all designated States
71-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

72	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
72-1	page	100
72-2	line	13
72-3	Identification of Deposit	
72-3-1	Name of depositary institution	American Type Culture Collection
72-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
72-3-3	Date of deposit	23 March 1999 (23.03.1999)
72-3-4	Accession Number	ATCC 203861
72-4	Additional Indications	NONE
72-5	Designated States for Which Indications are Made	all designated States
72-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
73	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
73-1	page	100
73-2	line	14
73-3	Identification of Deposit	
73-3-1	Name of depositary institution	American Type Culture Collection
73-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
73-3-3	Date of deposit	16 March 1999 (16.03.1999)
73-3-4	Accession Number	ATCC 203844
73-4	Additional Indications	NONE
73-5	Designated States for Which Indications are Made	all designated States
73-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
74	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
74-1	page	100
74-2	line	15
74-3	Identification of Deposit	
74-3-1	Name of depositary institution	American Type Culture Collection
74-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
74-3-3	Date of deposit	10 August 1999 (10.08.1999)
74-3-4	Accession Number	ATCC PTA-513
74-4	Additional Indications	NONE
74-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

74-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
75	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
75-1	page	100
75-2	line	16
75-3	Identification of Deposit	
75-3-1	Name of depositary institution	American Type Culture Collection
75-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
75-3-3	Date of deposit	09 February 1999 (09.02.1999)
75-3-4	Accession Number	ATCC 203663
75-4	Additional Indications	NONE
75-5	Designated States for Which Indications are Made	all designated States
76-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
76	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
76-1	page	100
76-2	line	17
76-3	Identification of Deposit	
76-3-1	Name of depositary institution	American Type Culture Collection
76-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
76-3-3	Date of deposit	16 March 1999 (16.03.1999)
76-3-4	Accession Number	ATCC 203851
76-4	Additional Indications	NONE
76-5	Designated States for Which Indications are Made	all designated States
76-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
77	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
77-1	page	100
77-2	line	18
77-3	Identification of Deposit	
77-3-1	Name of depositary institution	American Type Culture Collection
77-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
77-3-3	Date of deposit	20 April 1999 (20.04.1999)
77-3-4	Accession Number	ATCC 203950

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

77-4	Additional Indications	NONE
77-5	Designated States for Which Indications are Made	all designated States
77-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
78	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
78-1	page	100
78-2	line	19
78-3	Identification of Deposit	
78-3-1	Name of depositary institution	American Type Culture Collection
78-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
78-3-3	Date of deposit	30 March 1999 (30.03.1999)
78-3-4	Accession Number	ATCC 203895
78-4	Additional Indications	NONE
78-5	Designated States for Which Indications are Made	all designated States
78-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
79	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
79-1	page	100
79-2	line	20
79-3	Identification of Deposit	
79-3-1	Name of depositary institution	American Type Culture Collection
79-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
79-3-3	Date of deposit	25 May 1999 (25.05.1999)
79-3-4	Accession Number	ATCC PTA-134
79-4	Additional Indications	NONE
79-5	Designated States for Which Indications are Made	all designated States
79-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
80	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
80-1	page	100
80-2	line	21

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

80-3	Identification of Deposit	
80-3-1	Name of depositary institution	American Type Culture Collection
80-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
80-3-3	Date of deposit	16 March 1999 (16.03.1999)
80-3-4	Accession Number	ATCC 203852
80-4	Additional Indications	NONE
80-5	Designated States for Which Indications are Made	all designated States
80-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
81	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
81-1	page	100
81-2	line	22
81-3	Identification of Deposit	
81-3-1	Name of depositary institution	American Type Culture Collection
81-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
81-3-3	Date of deposit	22 June 1999 (22.06.1999)
81-3-4	Accession Number	ATCC PTA-258
81-4	Additional Indications	NONE
81-5	Designated States for Which Indications are Made	all designated States
81-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
82	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
82-1	page	100
82-2	line	23
82-3	Identification of Deposit	
82-3-1	Name of depositary institution	American Type Culture Collection
82-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
82-3-3	Date of deposit	22 June 1999 (22.06.1999)
82-3-4	Accession Number	ATCC PTA-259
82-4	Additional Indications	NONE
82-5	Designated States for Which Indications are Made	all designated States
82-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

83	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
83-1	page	100
83-2	line	24
83-3	Identification of Deposit	
83-3-1	Name of depositary institution	American Type Culture Collection
83-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
83-3-3	Date of deposit	23 March 1999 (23.03.1999)
83-3-4	Accession Number	ATCC 203866
83-4	Additional Indications	NONE
83-5	Designated States for Which Indications are Made	all designated States
83-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
84	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
84-1	page	100
84-2	line	25
84-3	Identification of Deposit	
84-3-1	Name of depositary institution	American Type Culture Collection
84-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
84-3-3	Date of deposit	16 March 1999 (16.03.1999)
84-3-4	Accession Number	ATCC 203853
84-4	Additional Indications	NONE
84-5	Designated States for Which Indications are Made	all designated States
84-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
85	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
85-1	page	100
85-2	line	26
85-3	Identification of Deposit	
85-3-1	Name of depositary institution	American Type Culture Collection
85-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
85-3-3	Date of deposit	30 March 1999 (30.03.1999)
85-3-4	Accession Number	ATCC 203892
85-4	Additional Indications	NONE
85-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

85-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
86	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
86-1	page	100
86-2	line	27
86-3	Identification of Deposit	
86-3-1	Name of depositary institution	American Type Culture Collection
86-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
86-3-3	Date of deposit	16 March 1999 (16.03.1999)
86-3-4	Accession Number	ATCC 203847
86-4	Additional Indications	NONE
86-5	Designated States for Which Indications are Made	all designated States
86-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
87	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
87-1	page	100
87-2	line	28
87-3	Identification of Deposit	
87-3-1	Name of depositary institution	American Type Culture Collection
87-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
87-3-3	Date of deposit	04 May 1999 (04.05.1999)
87-3-4	Accession Number	ATCC PTA-21
87-4	Additional Indications	NONE
87-5	Designated States for Which Indications are Made	all designated States
87-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
88	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
88-1	page	100
88-2	line	29
88-3	Identification of Deposit	
88-3-1	Name of depositary institution	American Type Culture Collection
88-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
88-3-3	Date of deposit	25 May 1999 (25.05.1999)
88-3-4	Accession Number	ATCC PTA-121

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

88-4	Additional Indications	NONE
88-5	Designated States for Which Indications are Made	all designated States
88-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
89	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
89-1	page	100
89-2	line	30
89-3	Identification of Deposit	
89-3-1	Name of depositary institution	American Type Culture Collection
89-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
89-3-3	Date of deposit	20 April 1999 (20.04.1999)
89-3-4	Accession Number	ATCC 203951
89-4	Additional Indications	NONE
89-5	Designated States for Which Indications are Made	all designated States
89-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
90	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
90-1	page	100
90-2	line	31
90-3	Identification of Deposit	
90-3-1	Name of depositary institution	American Type Culture Collection
90-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
90-3-3	Date of deposit	23 March 1999 (23.03.1999)
90-3-4	Accession Number	ATCC 203869
90-4	Additional Indications	NONE
90-5	Designated States for Which Indications are Made	all designated States
90-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
91	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
91-1	page	100
91-2	line	32

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

91-3	Identification of Deposit	
91-3-1	Name of depositary institution	American Type Culture Collection
91-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
91-3-3	Date of deposit	15 June 1999 (15.06.1999)
91-3-4	Accession Number	ATCC PTA-232
91-4	Additional Indications	NONE
91-5	Designated States for Which Indications are Made	all designated States
91-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
92	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
92-1	page	100
92-2	line	33
92-3	Identification of Deposit	
92-3-1	Name of depositary institution	American Type Culture Collection
92-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
92-3-3	Date of deposit	20 July 1999 (20.07.1999)
92-3-4	Accession Number	ATCC PTA-385
92-4	Additional Indications	NONE
92-5	Designated States for Which Indications are Made	all designated States
92-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
93	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
93-1	page	100
93-2	line	34
93-3	Identification of Deposit	
93-3-1	Name of depositary institution	American Type Culture Collection
93-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
93-3-3	Date of deposit	23 March 1999 (23.03.1999)
93-3-4	Accession Number	ATCC 203864
93-4	Additional Indications	NONE
93-5	Designated States for Which Indications are Made	all designated States
93-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

94	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
94-1	page	100
94-2	line	35
94-3	Identification of Deposit	
94-3-1	Name of depositary institution	American Type Culture Collection
94-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
94-3-3	Date of deposit	22 June 1999 (22.06.1999)
94-3-4	Accession Number	ATCC PTA-262
94-4	Additional Indications	NONE
94-5	Designated States for Which Indications are Made	all designated States
94-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
95	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
95-1	page	100
95-2	line	36
95-3	Identification of Deposit	
95-3-1	Name of depositary institution	American Type Culture Collection
95-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
95-3-3	Date of deposit	20 July 1999 (20.07.1999)
95-3-4	Accession Number	ATCC PTA-381
95-4	Additional Indications	NONE
95-5	Designated States for Which Indications are Made	all designated States
95-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
96	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
96-1	page	100
96-2	line	37
96-3	Identification of Deposit	
96-3-1	Name of depositary institution	American Type Culture Collection
96-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
96-3-3	Date of deposit	04 May 1999 (04.05.1999)
96-3-4	Accession Number	ATCC PTA-15
96-4	Additional Indications	NONE
96-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

96-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
97	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
97-1	page	100
97-2	line	38
97-3	Identification of Deposit	
97-3-1	Name of depositary institution	American Type Culture Collection
97-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
97-3-3	Date of deposit	15 June 1999 (15.06.1999)
97-3-4	Accession Number	ATCC PTA-239
97-4	Additional Indications	NONE
97-5	Designated States for Which Indications are Made	all designated States
97-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
98	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
98-1	page	100
98-2	line	39
98-3	Identification of Deposit	
98-3-1	Name of depositary institution	American Type Culture Collection
98-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
98-3-3	Date of deposit	20 July 1999 (20.07.1999)
98-3-4	Accession Number	ATCC PTA-384
98-4	Additional Indications	NONE
98-5	Designated States for Which Indications are Made	all designated States
98-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
99	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
99-1	page	100
99-2	line	40
99-3	Identification of Deposit	
99-3-1	Name of depositary institution	American Type Culture Collection
99-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
99-3-3	Date of deposit	03 August 1999 (03.08.1999)
99-3-4	Accession Number	ATCC PTA-475

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

99-4	Additional Indications	NONE
99-5	Designated States for Which Indications are Made	all designated States
99-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
100	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
100-1	page	100
100-2	line	41
100-3	Identification of Deposit	
100-3-1	Name of depositary institution	American Type Culture Collection
100-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
100-3-3	Date of deposit	16 March 1999 (16.03.1999)
100-3-4	Accession Number	ATCC 203854
100-4	Additional Indications	NONE
100-5	Designated States for Which Indications are Made	all designated States
100-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
101	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
101-1	page	100
101-2	line	42
101-3	Identification of Deposit	
101-3-1	Name of depositary institution	American Type Culture Collection
101-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
101-3-3	Date of deposit	20 July 1999 (20.07.1999)
101-3-4	Accession Number	ATCC PTA-378
101-4	Additional Indications	NONE
101-5	Designated States for Which Indications are Made	all designated States
101-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
102	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
102-1	page	100
102-2	line	43

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

102-3	Identification of Deposit	
102-3-1	Name of depositary institution	American Type Culture Collection
102-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
102-3-3	Date of deposit	22 June 1999 (22.06.1999)
102-3-4	Accession Number	ATCC PTA-257
102-4	Additional Indications	NONE
102-5	Designated States for Which Indications are Made	all designated States
102-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
103	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
103-1	page	100
103-2	line	44
103-3	Identification of Deposit	
103-3-1	Name of depositary institution	American Type Culture Collection
103-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
103-3-3	Date of deposit	15 June 1999 (15.06.1999)
103-3-4	Accession Number	ATCC PTA-231
103-4	Additional Indications	NONE
103-5	Designated States for Which Indications are Made	all designated States
103-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
104	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
104-1	page	100
104-2	line	45
104-3	Identification of Deposit	
104-3-1	Name of depositary institution	American Type Culture Collection
104-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
104-3-3	Date of deposit	20 July 1999 (20.07.1999)
104-3-4	Accession Number	ATCC PTA-388
104-4	Additional Indications	NONE
104-5	Designated States for Which Indications are Made	all designated States
104-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

105	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
105-1	page	100
105-2	line	46
105-3	Identification of Deposit	
105-3-1	Name of depositary institution	American Type Culture Collection
105-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
105-3-3	Date of deposit	31 August 1999 (31.08.1999)
105-3-4	Accession Number	ATCC PTA-620
105-4	Additional Indications	NONE
105-5	Designated States for Which Indications are Made	all designated States
105-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
106	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
106-1	page	100
106-2	line	47
106-3	Identification of Deposit	
106-3-1	Name of depositary institution	American Type Culture Collection
106-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
106-3-3	Date of deposit	25 May 1999 (25.05.1999)
106-3-4	Accession Number	ATCC PTA-118
106-4	Additional Indications	NONE
106-5	Designated States for Which Indications are Made	all designated States
106-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
107	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
107-1	page	100
107-2	line	48
107-3	Identification of Deposit	
107-3-1	Name of depositary institution	American Type Culture Collection
107-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
107-3-3	Date of deposit	03 August 1999 (03.08.1999)
107-3-4	Accession Number	ATCC PTA-477
107-4	Additional Indications	NONE
107-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

107-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
108	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
108-1	page	100
108-2	line	49
108-3	Identification of Deposit	
108-3-1	Name of depositary institution	American Type Culture Collection
108-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
108-3-3	Date of deposit	03 August 1999 (03.08.1999)
108-3-4	Accession Number	ATCC PTA-488
108-4	Additional Indications	NONE
108-5	Designated States for Which Indications are Made	all designated States
108-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
109	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
109-1	page	100
109-2	line	50
109-3	Identification of Deposit	
109-3-1	Name of depositary institution	American Type Culture Collection
109-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
109-3-3	Date of deposit	16 March 1999 (16.03.1999)
109-3-4	Accession Number	ATCC 203849
109-4	Additional Indications	NONE
109-5	Designated States for Which Indications are Made	all designated States
109-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
110	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
110-1	page	100
110-2	line	51
110-3	Identification of Deposit	
110-3-1	Name of depositary institution	American Type Culture Collection
110-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
110-3-3	Date of deposit	09 March 1999 (09.03.1999)
110-3-4	Accession Number	ATCC 203837

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

110-4	Additional Indications	NONE
110-5	Designated States for Which Indications are Made	all designated States
110-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
111	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
111-1	page	100
111-2	line	52
111-3	Identification of Deposit	
111-3-1	Name of depositary institution	American Type Culture Collection
111-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
111-3-3	Date of deposit	20 July 1999 (20.07.1999)
111-3-4	Accession Number	ATCC PTA-380
111-4	Additional Indications	NONE
111-5	Designated States for Which Indications are Made	all designated States
111-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
112	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
112-1	page	100
112-2	line	53
112-3	Identification of Deposit	
112-3-1	Name of depositary institution	American Type Culture Collection
112-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
112-3-3	Date of deposit	11 May 1999 (11.05.1999)
112-3-4	Accession Number	ATCC PTA-44
112-4	Additional Indications	NONE
112-5	Designated States for Which Indications are Made	all designated States
112-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
113	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
113-1	page	100
113-2	line	54

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

113-3	Identification of Deposit	
113-3-1	Name of depositary institution	American Type Culture Collection
113-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
113-3-3	Date of deposit	11 May 1999 (11.05.1999)
113-3-4	Accession Number	ATCC PTA-42
113-4	Additional Indications	NONE
113-5	Designated States for Which Indications are Made	all designated States
113-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
114	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
114-1	page	100
114-2	line	55
114-3	Identification of Deposit	
114-3-1	Name of depositary institution	American Type Culture Collection
114-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
114-3-3	Date of deposit	25 May 1999 (25.05.1999)
114-3-4	Accession Number	ATCC PTA-123
114-4	Additional Indications	NONE
114-5	Designated States for Which Indications are Made	all designated States
114-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
115	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
115-1	page	101
115-2	line	2
115-3	Identification of Deposit	
115-3-1	Name of depositary institution	American Type Culture Collection
115-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
115-3-3	Date of deposit	03 August 1999 (03.08.1999)
115-3-4	Accession Number	ATCC PTA-482
115-4	Additional Indications	NONE
115-5	Designated States for Which Indications are Made	all designated States
115-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

116	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
116-1	page	101
116-2	line	3
116-3	Identification of Deposit	
116-3-1	Name of depositary institution	American Type Culture Collection
116-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
116-3-3	Date of deposit	03 August 1999 (03.08.1999)
116-3-4	Accession Number	ATCC PTA-483
116-4	Additional Indications	NONE
116-5	Designated States for Which Indications are Made	all designated States
116-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
117	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
117-1	page	101
117-2	line	4
117-3	Identification of Deposit	
117-3-1	Name of depositary institution	American Type Culture Collection
117-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
117-3-3	Date of deposit	03 August 1999 (03.08.1999)
117-3-4	Accession Number	ATCC PTA-485
117-4	Additional Indications	NONE
117-5	Designated States for Which Indications are Made	all designated States
117-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
118	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
118-1	page	101
118-2	line	5
118-3	Identification of Deposit	
118-3-1	Name of depositary institution	American Type Culture Collection
118-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
118-3-3	Date of deposit	03 August 1999 (03.08.1999)
118-3-4	Accession Number	ATCC PTA-480
118-4	Additional Indications	NONE
118-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

118-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
119	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
119-1	page	101
119-2	line	6
119-3	Identification of Deposit	
119-3-1	Name of depositary institution	American Type Culture Collection
119-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
119-3-3	Date of deposit	03 August 1999 (03.08.1999)
119-3-4	Accession Number	ATCC PTA-476
119-4	Additional Indications	NONE
119-5	Designated States for Which Indications are Made	all designated States
119-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
120	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
120-1	page	101
120-2	line	7
120-3	Identification of Deposit	
120-3-1	Name of depositary institution	American Type Culture Collection
120-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
120-3-3	Date of deposit	03 August 1999 (03.08.1999)
120-3-4	Accession Number	ATCC PTA-472
120-4	Additional Indications	NONE
120-5	Designated States for Which Indications are Made	all designated States
120-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
121	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
121-1	page	101
121-2	line	8
121-3	Identification of Deposit	
121-3-1	Name of depositary institution	American Type Culture Collection
121-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
121-3-3	Date of deposit	03 August 1999 (03.08.1999)
121-3-4	Accession Number	ATCC PTA-487

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

121-4	Additional Indications	NONE
121-5	Designated States for Which Indications are Made	all designated States
121-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
122	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
122-1	page	101
122-2	line	9
122-3	Identification of Deposit	
122-3-1	Name of depositary institution	American Type Culture Collection
122-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
122-3-3	Date of deposit	03 August 1999 (03.08.1999)
122-3-4	Accession Number	ATCC PTA-484
122-4	Additional Indications	NONE
122-5	Designated States for Which Indications are Made	all designated States
122-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
123	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
123-1	page	101
123-2	line	10
123-3	Identification of Deposit	
123-3-1	Name of depositary institution	American Type Culture Collection
123-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
123-3-3	Date of deposit	17 August 1999 (17.08.1999)
123-3-4	Accession Number	ATCC PTA-546
123-4	Additional Indications	NONE
123-5	Designated States for Which Indications are Made	all designated States
123-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
124	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
124-1	page	101
124-2	line	11

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

124-3	Identification of Deposit	
124-3-1	Name of depositary institution	American Type Culture Collection
124-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
124-3-3	Date of deposit	10 August 1999 (10.08.1999)
124-3-4	Accession Number	ATCC PTA-515
124-4	Additional Indications	NONE
124-5	Designated States for Which Indications are Made	all designated States
124-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
125	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
125-1	page	101
125-2	line	12
125-3	Identification of Deposit	
125-3-1	Name of depositary institution	American Type Culture Collection
125-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
125-3-3	Date of deposit	19 October 1999 (19.10.1999)
125-3-4	Accession Number	ATCC PTA-861
125-4	Additional Indications	NONE
125-5	Designated States for Which Indications are Made	all designated States
125-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
126	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
126-1	page	101
126-2	line	13
126-3	Identification of Deposit	
126-3-1	Name of depositary institution	American Type Culture Collection
126-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
126-3-3	Date of deposit	10 August 1999 (10.08.1999)
126-3-4	Accession Number	ATCC PTA-518
126-4	Additional Indications	NONE
126-5	Designated States for Which Indications are Made	all designated States
126-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

127	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
127-1	page	101
127-2	line	14
127-3	Identification of Deposit	
127-3-1	Name of depositary institution	American Type Culture Collection
127-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
127-3-3	Date of deposit	10 August 1999 (10.08.1999)
127-3-4	Accession Number	ATCC PTA-512
127-4	Additional Indications	NONE
127-5	Designated States for Which Indications are Made	all designated States
127-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
128	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
128-1	page	101
128-2	line	15
128-3	Identification of Deposit	
128-3-1	Name of depositary institution	American Type Culture Collection
128-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
128-3-3	Date of deposit	03 August 1999 (03.08.1999)
128-3-4	Accession Number	ATCC PTA-489
128-4	Additional Indications	NONE
128-5	Designated States for Which Indications are Made	all designated States
128-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
129	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
129-1	page	101
129-2	line	16
129-3	Identification of Deposit	
129-3-1	Name of depositary institution	American Type Culture Collection
129-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
129-3-3	Date of deposit	31 August 1999 (31.08.1999)
129-3-4	Accession Number	ATCC PTA-614
129-4	Additional Indications	NONE
129-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

129-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
130	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
130-1	page	101
130-2	line	17
130-3	Identification of Deposit	
130-3-1	Name of depositary institution	American Type Culture Collection
130-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
130-3-3	Date of deposit	16 November 1999 (16.11.1999)
130-3-4	Accession Number	ATCC PTA-957
130-4	Additional Indications	NONE
130-5	Designated States for Which Indications are Made	all designated States
130-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
131	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
131-1	page	101
131-2	line	18
131-3	Identification of Deposit	
131-3-1	Name of depositary institution	American Type Culture Collection
131-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
131-3-3	Date of deposit	05 October 1999 (05.10.1999)
131-3-4	Accession Number	ATCC PTA-819
131-4	Additional Indications	NONE
131-5	Designated States for Which Indications are Made	all designated States
131-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
132	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
132-1	page	101
132-2	line	19
132-3	Identification of Deposit	
132-3-1	Name of depositary institution	American Type Culture Collection
132-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
132-3-3	Date of deposit	18 September 1997 (18.09.1997)
132-3-4	Accession Number	ATCC 209280

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

132-4	Additional Indications	NONE
132-5	Designated States for Which Indications are Made	all designated States
132-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
133	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
133-1	page	101
133-2	line	20
133-3	Identification of Deposit	
133-3-1	Name of depositary institution	American Type Culture Collection
133-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
133-3-3	Date of deposit	14 April 1998 (14.04.1998)
133-3-4	Accession Number	ATCC 209772
133-4	Additional Indications	NONE
133-5	Designated States for Which Indications are Made	all designated States
133-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
134	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
134-1	page	101
134-2	line	21
134-3	Identification of Deposit	
134-3-1	Name of depositary institution	American Type Culture Collection
134-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
134-3-3	Date of deposit	16 October 1997 (16.10.1997)
134-3-4	Accession Number	ATCC 209375
134-4	Additional Indications	NONE
134-5	Designated States for Which Indications are Made	all designated States
134-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
135	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
135-1	page	101
135-2	line	22

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

135-3	Identification of Deposit	
135-3-1	Name of depositary institution	American Type Culture Collection
135-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
135-3-3	Date of deposit	23 September 1997 (23.09.1997)
135-3-4	Accession Number	ATCC 209296
135-4	Additional Indications	NONE
135-5	Designated States for Which Indications are Made	all designated States
135-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
136	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
136-1	page	101
136-2	line	23
136-3	Identification of Deposit	
136-3-1	Name of depositary institution	American Type Culture Collection
136-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
136-3-3	Date of deposit	18 September 1997 (18.09.1997)
136-3-4	Accession Number	ATCC 209279
136-4	Additional Indications	NONE
136-5	Designated States for Which Indications are Made	all designated States
136-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
137	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
137-1	page	101
137-2	line	24
137-3	Identification of Deposit	
137-3-1	Name of depositary institution	American Type Culture Collection
137-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
137-3-3	Date of deposit	05 March 1998 (05.03.1998)
137-3-4	Accession Number	ATCC 209653
137-4	Additional Indications	NONE
137-5	Designated States for Which Indications are Made	all designated States
137-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

138	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
138-1	page	101
138-2	line	25
138-3	Identification of Deposit	
138-3-1	Name of depositary institution	American Type Culture Collection
138-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
138-3-3	Date of deposit	16 October 1997 (16.10.1997)
138-3-4	Accession Number	ATCC 209385
138-4	Additional Indications	NONE
138-5	Designated States for Which Indications are Made	all designated States
138-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
139	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
139-1	page	101
139-2	line	26
139-3	Identification of Deposit	
139-3-1	Name of depositary institution	American Type Culture Collection
139-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
139-3-3	Date of deposit	16 September 1997 (16.09.1997)
139-3-4	Accession Number	ATCC 209261
139-4	Additional Indications	NONE
139-5	Designated States for Which Indications are Made	all designated States
139-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
140	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
140-1	page	101
140-2	line	27
140-3	Identification of Deposit	
140-3-1	Name of depositary institution	American Type Culture Collection
140-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
140-3-3	Date of deposit	16 October 1997 (16.10.1997)
140-3-4	Accession Number	ATCC 209384
140-4	Additional Indications	NONE
140-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

140-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
141	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
141-1	page	101
141-2	line	28
141-3	Identification of Deposit	
141-3-1	Name of depositary institution	American Type Culture Collection
141-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
141-3-3	Date of deposit	16 September 1997 (16.09.1997)
141-3-4	Accession Number	ATCC 209258
141-4	Additional Indications	NONE
141-5	Designated States for Which Indications are Made	all designated States
141-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
142	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
142-1	page	101
142-2	line	29
142-3	Identification of Deposit	
142-3-1	Name of depositary institution	American Type Culture Collection
142-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
142-3-3	Date of deposit	16 September 1997 (16.09.1997)
142-3-4	Accession Number	ATCC 209257
142-4	Additional Indications	NONE
142-5	Designated States for Which Indications are Made	all designated States
142-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
143	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
143-1	page	101
143-2	line	30
143-3	Identification of Deposit	
143-3-1	Name of depositary institution	American Type Culture Collection
143-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
143-3-3	Date of deposit	30 May 1997 (30.05.1997)
143-3-4	Accession Number	ATCC 209087

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

143-4	Additional Indications	NONE
143-5	Designated States for Which Indications are Made	all designated States
143-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
144	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
144-1	page	101
144-2	line	31
144-3	Identification of Deposit	
144-3-1	Name of depositary institution	American Type Culture Collection
144-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
144-3-3	Date of deposit	16 October 1997 (16.10.1997)
144-3-4	Accession Number	ATCC 209381
144-4	Additional Indications	NONE
144-5	Designated States for Which Indications are Made	all designated States
144-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
145	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
145-1	page	101
145-2	line	32
145-3	Identification of Deposit	
145-3-1	Name of depositary institution	American Type Culture Collection
145-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
145-3-3	Date of deposit	16 September 1997 (16.09.1997)
145-3-4	Accession Number	ATCC 209262
145-4	Additional Indications	NONE
145-5	Designated States for Which Indications are Made	all designated States
145-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
146	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
146-1	page	101
146-2	line	33

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

146-3	Identification of Deposit	
146-3-1	Name of depositary institution	American Type Culture Collection
146-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
146-3-3	Date of deposit	28 October 1997 (28.10.1997)
146-3-4	Accession Number	ATCC 209420
146-4	Additional Indications	NONE
146-5	Designated States for Which Indications are Made	all designated States
146-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
147	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
147-1	page	101
147-2	line	34
147-3	Identification of Deposit	
147-3-1	Name of depositary institution	American Type Culture Collection
147-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
147-3-3	Date of deposit	16 September 1997 (16.09.1997)
147-3-4	Accession Number	ATCC 209256
147-4	Additional Indications	NONE
147-5	Designated States for Which Indications are Made	all designated States
147-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
148	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
148-1	page	101
148-2	line	35
148-3	Identification of Deposit	
148-3-1	Name of depositary institution	American Type Culture Collection
148-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
148-3-3	Date of deposit	16 September 1997 (16.09.1997)
148-3-4	Accession Number	ATCC 209251
148-4	Additional Indications	NONE
148-5	Designated States for Which Indications are Made	all designated States
148-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

149	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
149-1	page	101
149-2	line	36
149-3	Identification of Deposit	
149-3-1	Name of depositary institution	American Type Culture Collection
149-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
149-3-3	Date of deposit	16 September 1997 (16.09.1997)
149-3-4	Accession Number	ATCC 209263
149-4	Additional Indications	NONE
149-5	Designated States for Which Indications are Made	all designated States
149-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
150	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
150-1	page	101
150-2	line	37
150-3	Identification of Deposit	
150-3-1	Name of depositary institution	American Type Culture Collection
150-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
150-3-3	Date of deposit	16 September 1997 (16.09.1997)
150-3-4	Accession Number	ATCC 209264
150-4	Additional Indications	NONE
150-5	Designated States for Which Indications are Made	all designated States
150-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
151	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
151-1	page	101
151-2	line	38
151-3	Identification of Deposit	
151-3-1	Name of depositary institution	American Type Culture Collection
151-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
151-3-3	Date of deposit	16 October 1997 (16.10.1997)
151-3-4	Accession Number	ATCC 209376
151-4	Additional Indications	NONE
151-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

151-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
152	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
152-1	page	101
152-2	line	39
152-3	Identification of Deposit	
152-3-1	Name of depositary institution	American Type Culture Collection
152-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
152-3-3	Date of deposit	17 October 1997 (17.10.1997)
152-3-4	Accession Number	ATCC 209391
152-4	Additional Indications	NONE
152-5	Designated States for Which Indications are Made	all designated States
152-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
153	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
153-1	page	101
153-2	line	40
153-3	Identification of Deposit	
153-3-1	Name of depositary institution	American Type Culture Collection
153-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
153-3-3	Date of deposit	28 October 1997 (28.10.1997)
153-3-4	Accession Number	ATCC 209417
153-4	Additional Indications	NONE
153-5	Designated States for Which Indications are Made	all designated States
153-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
154	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
154-1	page	101
154-2	line	41
154-3	Identification of Deposit	
154-3-1	Name of depositary institution	American Type Culture Collection
154-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
154-3-3	Date of deposit	16 September 1997 (16.09.1997)
154-3-4	Accession Number	ATCC 209253

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

154-4	Additional Indications	NONE
154-5	Designated States for Which Indications are Made	all designated States
154-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
155	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
155-1	page	101
155-2	line	42
155-3	Identification of Deposit	
155-3-1	Name of depositary institution	American Type Culture Collection
155-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
155-3-3	Date of deposit	12 May 1998 (12.05.1998)
155-3-4	Accession Number	ATCC 209855
155-4	Additional Indications	NONE
155-5	Designated States for Which Indications are Made	all designated States
155-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
156	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
156-1	page	101
156-2	line	43
156-3	Identification of Deposit	
156-3-1	Name of depositary institution	American Type Culture Collection
156-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
156-3-3	Date of deposit	10 December 1997 (10.12.1997)
156-3-4	Accession Number	ATCC 209526
156-4	Additional Indications	NONE
156-5	Designated States for Which Indications are Made	all designated States
156-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
157	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
157-1	page	101
157-2	line	44

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

157-3	Identification of Deposit	
157-3-1	Name of depositary institution	American Type Culture Collection
157-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
157-3-3	Date of deposit	16 September 1997 (16.09.1997)
157-3-4	Accession Number	ATCC 209252
157-4	Additional Indications	NONE
157-5	Designated States for Which Indications are Made	all designated States
157-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
158	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
158-1	page	101
158-2	line	45
158-3	Identification of Deposit	
158-3-1	Name of depositary institution	American Type Culture Collection
158-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
158-3-3	Date of deposit	16 October 1997 (16.10.1997)
158-3-4	Accession Number	ATCC 209374
158-4	Additional Indications	NONE
158-5	Designated States for Which Indications are Made	all designated States
158-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
159	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
159-1	page	101
159-2	line	46
159-3	Identification of Deposit	
159-3-1	Name of depositary institution	American Type Culture Collection
159-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
159-3-3	Date of deposit	10 December 1997 (10.12.1997)
159-3-4	Accession Number	ATCC 209528
159-4	Additional Indications	NONE
159-5	Designated States for Which Indications are Made	all designated States
159-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

160	The Indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
160-1	page	101
160-2	line	47
160-3	Identification of Deposit	
160-3-1	Name of depositary institution	American Type Culture Collection
160-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
160-3-3	Date of deposit	16 September 1997 (16.09.1997)
160-3-4	Accession Number	ATCC 209265
160-4	Additional Indications	NONE
160-5	Designated States for Which Indications are Made	all designated States
160-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
161	The Indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
161-1	page	101
161-2	line	48
161-3	Identification of Deposit	
161-3-1	Name of depositary institution	American Type Culture Collection
161-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
161-3-3	Date of deposit	17 October 1997 (17.10.1997)
161-3-4	Accession Number	ATCC 209396
161-4	Additional Indications	NONE
161-5	Designated States for Which Indications are Made	all designated States
161-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
162	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
162-1	page	101
162-2	line	49
162-3	Identification of Deposit	
162-3-1	Name of depositary institution	American Type Culture Collection
162-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
162-3-3	Date of deposit	18 August 1997 (18.08.1997)
162-3-4	Accession Number	ATCC 209201
162-4	Additional Indications	NONE
162-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

162-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
163	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
163-1	page	101
163-2	line	50
163-3	Identification of Deposit	
163-3-1	Name of depositary institution	American Type Culture Collection
163-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
163-3-3	Date of deposit	28 October 1997 (28.10.1997)
163-3-4	Accession Number	ATCC 209416
163-4	Additional Indications	NONE
163-5	Designated States for Which Indications are Made	all designated States
163-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
164	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
164-1	page	101
164-2	line	51
164-3	Identification of Deposit	
164-3-1	Name of depositary institution	American Type Culture Collection
164-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
164-3-3	Date of deposit	17 October 1997 (17.10.1997)
164-3-4	Accession Number	ATCC 209403
164-4	Additional Indications	NONE
164-5	Designated States for Which Indications are Made	all designated States
164-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
165	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
165-1	page	101
165-2	line	52
165-3	Identification of Deposit	
165-3-1	Name of depositary institution	American Type Culture Collection
165-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
165-3-3	Date of deposit	28 October 1997 (28.10.1997)
165-3-4	Accession Number	ATCC 209419

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

165-4	Additional Indications	NONE
165-5	Designated States for Which Indications are Made	all designated States
165-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
166	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
166-1	page	101
166-2	line	53
166-3	Identification of Deposit	
166-3-1	Name of depositary institution	American Type Culture Collection
166-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
166-3-3	Date of deposit	17 October 1997 (17.10.1997)
166-3-4	Accession Number	ATCC 209402
166-4	Additional Indications	NONE
166-5	Designated States for Which Indications are Made	all designated States
166-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
167	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
167-1	page	101
167-2	line	54
167-3	Identification of Deposit	
167-3-1	Name of depositary institution	American Type Culture Collection
167-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
167-3-3	Date of deposit	16 October 1997 (16.10.1997)
167-3-4	Accession Number	ATCC 209378
167-4	Additional Indications	NONE
167-5	Designated States for Which Indications are Made	all designated States
167-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
168	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
168-1	page	101
168-2	line	55

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

168-3	Identification of Deposit	
168-3-1	Name of depositary institution	American Type Culture Collection
168-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
168-3-3	Date of deposit	21 November 1997 (21.11.1997)
168-3-4	Accession Number	ATCC 209489
168-4	Additional Indications	NONE
168-5	Designated States for Which Indications are Made	all designated States
168-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
169	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
169-1	page	102
169-2	line	2
169-3	Identification of Deposit	
169-3-1	Name of depositary institution	American Type Culture Collection
169-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
169-3-3	Date of deposit	17 October 1997 (17.10.1997)
169-3-4	Accession Number	ATCC 209401
169-4	Additional Indications	NONE
169-5	Designated States for Which Indications are Made	all designated States
169-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
170	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
170-1	page	102
170-2	line	3
170-3	Identification of Deposit	
170-3-1	Name of depositary institution	American Type Culture Collection
170-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
170-3-3	Date of deposit	17 October 1997 (17.10.1997)
170-3-4	Accession Number	ATCC 209397
170-4	Additional Indications	NONE
170-5	Designated States for Which Indications are Made	all designated States
170-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

171	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
171-1	page	102
171-2	line	4
171-3	Identification of Deposit	
171-3-1	Name of depositary institution	American Type Culture Collection
171-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
171-3-3	Date of deposit	17 October 1997 (17.10.1997)
171-3-4	Accession Number	ATCC 209389
171-4	Additional Indications	NONE
171-5	Designated States for Which Indications are Made	all designated States
171-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
172	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
172-1	page	102
172-2	line	5
172-3	Identification of Deposit	
172-3-1	Name of depositary institution	American Type Culture Collection
172-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
172-3-3	Date of deposit	07 November 1997 (07.11.1997)
172-3-4	Accession Number	ATCC 209438
172-4	Additional Indications	NONE
172-5	Designated States for Which Indications are Made	all designated States
172-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
173	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
173-1	page	102
173-2	line	6
173-3	Identification of Deposit	
173-3-1	Name of depositary institution	American Type Culture Collection
173-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
173-3-3	Date of deposit	21 November 1997 (21.11.1997)
173-3-4	Accession Number	ATCC 209492
173-4	Additional Indications	NONE
173-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

173-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
174	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
174-1	page	102
174-2	line	7
174-3	Identification of Deposit	
174-3-1	Name of depositary institution	American Type Culture Collection
174-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
174-3-3	Date of deposit	17 October 1997 (17.10.1997)
174-3-4	Accession Number	ATCC 209388
174-4	Additional Indications	NONE
174-5	Designated States for Which Indications are Made	all designated States
174-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
175	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
175-1	page	102
175-2	line	8
175-3	Identification of Deposit	
175-3-1	Name of depositary institution	American Type Culture Collection
175-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
175-3-3	Date of deposit	07 November 1997 (07.11.1997)
175-3-4	Accession Number	ATCC 209432
175-4	Additional Indications	NONE
175-5	Designated States for Which Indications are Made	all designated States
175-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
176	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
176-1	page	102
176-2	line	9
176-3	Identification of Deposit	
176-3-1	Name of depositary institution	American Type Culture Collection
176-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
176-3-3	Date of deposit	07 November 1997 (07.11.1997)
176-3-4	Accession Number	ATCC 209439

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

176-4	Additional Indications	NONE
176-5	Designated States for Which Indications are Made	all designated States
176-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
177	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
177-1	page	102
177-2	line	10
177-3	Identification of Deposit	
177-3-1	Name of depositary institution	American Type Culture Collection
177-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
177-3-3	Date of deposit	07 November 1997 (07.11.1997)
177-3-4	Accession Number	ATCC 209433
177-4	Additional Indications	NONE
177-5	Designated States for Which Indications are Made	all designated States
177-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
178	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
178-1	page	102
178-2	line	11
178-3	Identification of Deposit	
178-3-1	Name of depositary institution	American Type Culture Collection
178-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
178-3-3	Date of deposit	05 February 1998 (05.02.1998)
178-3-4	Accession Number	ATCC 209618
178-4	Additional Indications	NONE
178-5	Designated States for Which Indications are Made	all designated States
178-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
179	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
179-1	page	102
179-2	line	12

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

179-3	Identification of Deposit	
179-3-1	Name of depositary institution	American Type Culture Collection
179-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
179-3-3	Date of deposit	21 November 1997 (21.11.1997)
179-3-4	Accession Number	ATCC 209484
179-4	Additional Indications	NONE
179-5	Designated States for Which Indications are Made	all designated States
179-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
180	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
180-1	page	102
180-2	line	13
180-3	Identification of Deposit	
180-3-1	Name of depositary institution	American Type Culture Collection
180-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
180-3-3	Date of deposit	21 November 1997 (21.11.1997)
180-3-4	Accession Number	ATCC 209487
180-4	Additional Indications	NONE
180-5	Designated States for Which Indications are Made	all designated States
180-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
181	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
181-1	page	102
181-2	line	14
181-3	Identification of Deposit	
181-3-1	Name of depositary institution	American Type Culture Collection
181-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
181-3-3	Date of deposit	07 November 1997 (07.11.1997)
181-3-4	Accession Number	ATCC 209434
181-4	Additional Indications	NONE
181-5	Designated States for Which Indications are Made	all designated States
181-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

182	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
182-1	page	102
182-2	line	15
182-3	Identification of Deposit	
182-3-1	Name of depositary institution	American Type Culture Collection
182-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
182-3-3	Date of deposit	26 March 1998 (26.03.1998)
182-3-4	Accession Number	ATCC 209704
182-4	Additional Indications	NONE
182-5	Designated States for Which Indications are Made	all designated States
182-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
183	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
183-1	page	102
183-2	line	16
183-3	Identification of Deposit	
183-3-1	Name of depositary institution	American Type Culture Collection
183-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
183-3-3	Date of deposit	28 April 1998 (28.04.1998)
183-3-4	Accession Number	ATCC 209808
183-4	Additional Indications	NONE
183-5	Designated States for Which Indications are Made	all designated States
183-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
184	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
184-1	page	102
184-2	line	17
184-3	Identification of Deposit	
184-3-1	Name of depositary institution	American Type Culture Collection
184-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
184-3-3	Date of deposit	06 May 1998 (06.05.1998)
184-3-4	Accession Number	ATCC 209847
184-4	Additional Indications	NONE
184-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

184-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
185	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
185-1	page	102
185-2	line	18
185-3	Identification of Deposit	
185-3-1	Name of depositary institution	American Type Culture Collection
185-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
185-3-3	Date of deposit	05 February 1998 (05.02.1998)
185-3-4	Accession Number	ATCC 209616
185-4	Additional Indications	NONE
185-5	Designated States for Which Indications are Made	all designated States
186-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
186	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
186-1	page	102
186-2	line	19
186-3	Identification of Deposit	
186-3-1	Name of depositary institution	American Type Culture Collection
186-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
186-3-3	Date of deposit	05 February 1998 (05.02.1998)
186-3-4	Accession Number	ATCC 209619
186-4	Additional Indications	NONE
186-5	Designated States for Which Indications are Made	all designated States
186-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
187	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
187-1	page	102
187-2	line	20
187-3	Identification of Deposit	
187-3-1	Name of depositary institution	American Type Culture Collection
187-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
187-3-3	Date of deposit	11 August 1998 (11.08.1998)
187-3-4	Accession Number	ATCC 203109

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

187-4	Additional Indications	NONE
187-5	Designated States for Which Indications are Made	all designated States
187-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
188	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
188-1	page	102
188-2	line	21
188-3	Identification of Deposit	
188-3-1	Name of depositary institution	American Type Culture Collection
188-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
188-3-3	Date of deposit	31 March 1998 (31.03.1998)
188-3-4	Accession Number	ATCC 209715
188-4	Additional Indications	NONE
188-5	Designated States for Which Indications are Made	all designated States
188-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
189	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
189-1	page	102
189-2	line	22
189-3	Identification of Deposit	
189-3-1	Name of depositary institution	American Type Culture Collection
189-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
189-3-3	Date of deposit	11 March 1998 (11.03.1998)
189-3-4	Accession Number	ATCC 209669
189-4	Additional Indications	NONE
189-5	Designated States for Which Indications are Made	all designated States
189-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
190	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
190-1	page	102
190-2	line	23

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

190-3	Identification of Deposit	
190-3-1	Name of depositary institution	American Type Culture Collection
190-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
190-3-3	Date of deposit	23 June 1998 (23.06.1998)
190-3-4	Accession Number	ATCC 203002
190-4	Additional Indications	NONE
190-5	Designated States for Which Indications are Made	all designated States
190-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
191	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
191-1	page	102
191-2	line	24
191-3	Identification of Deposit	
191-3-1	Name of depositary institution	American Type Culture Collection
191-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
191-3-3	Date of deposit	26 March 1998 (26.03.1998)
191-3-4	Accession Number	ATCC 209705
191-4	Additional Indications	NONE
191-5	Designated States for Which Indications are Made	all designated States
191-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
192	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
192-1	page	102
192-2	line	25
192-3	Identification of Deposit	
192-3-1	Name of depositary institution	American Type Culture Collection
192-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
192-3-3	Date of deposit	16 June 1998 (16.06.1998)
192-3-4	Accession Number	ATCC 209981
192-4	Additional Indications	NONE
192-5	Designated States for Which Indications are Made	all designated States
192-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

193	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
193-1	page	102
193-2	line	26
193-3	Identification of Deposit	
193-3-1	Name of depositary institution	American Type Culture Collection
193-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
193-3-3	Date of deposit	07 April 1998 (07.04.1998)
193-3-4	Accession Number	ATCC 209749
193-4	Additional Indications	NONE
193-5	Designated States for Which Indications are Made	all designated States
193-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
194	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
194-1	page	102
194-2	line	27
194-3	Identification of Deposit	
194-3-1	Name of depositary institution	American Type Culture Collection
194-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
194-3-3	Date of deposit	12 May 1998 (12.05.1998)
194-3-4	Accession Number	ATCC 209859
194-4	Additional Indications	NONE
194-5	Designated States for Which Indications are Made	all designated States
194-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
195	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
195-1	page	102
195-2	line	28
195-3	Identification of Deposit	
195-3-1	Name of depositary institution	American Type Culture Collection
195-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
195-3-3	Date of deposit	06 May 1998 (06.05.1998)
195-3-4	Accession Number	ATCC 209845
195-4	Additional Indications	NONE
195-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

195-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
196	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
196-1	page	102
196-2	line	29
196-3	Identification of Deposit	
196-3-1	Name of depositary institution	American Type Culture Collection
196-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
196-3-3	Date of deposit	07 April 1998 (07.04.1998)
196-3-4	Accession Number	ATCC 209748
196-4	Additional Indications	NONE
196-5	Designated States for Which Indications are Made	all designated States
196-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
197	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
197-1	page	102
197-2	line	30
197-3	Identification of Deposit	
197-3-1	Name of depositary institution	American Type Culture Collection
197-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
197-3-3	Date of deposit	11 August 1998 (11.08.1998)
197-3-4	Accession Number	ATCC 203107
197-4	Additional Indications	NONE
197-5	Designated States for Which Indications are Made	all designated States
197-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
198	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
198-1	page	102
198-2	line	31
198-3	Identification of Deposit	
198-3-1	Name of depositary institution	American Type Culture Collection
198-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
198-3-3	Date of deposit	23 April 1998 (23.04.1998)
198-3-4	Accession Number	ATCC 209801

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

198-4	Additional Indications	NONE
198-5	Designated States for Which Indications are Made	all designated States
198-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
199	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
199-1	page	102
199-2	line	32
199-3	Identification of Deposit	
199-3-1	Name of depositary institution	American Type Culture Collection
199-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
199-3-3	Date of deposit	09 June 1998 (09.06.1998)
199-3-4	Accession Number	ATCC 209948
199-4	Additional Indications	NONE
199-5	Designated States for Which Indications are Made	all designated States
199-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
200	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
200-1	page	102
200-2	line	33
200-3	Identification of Deposit	
200-3-1	Name of depositary institution	American Type Culture Collection
200-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
200-3-3	Date of deposit	20 May 1998 (20.05.1998)
200-3-4	Accession Number	ATCC 209883
200-4	Additional Indications	NONE
200-5	Designated States for Which Indications are Made	all designated States
200-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
201	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
201-1	page	102
201-2	line	34

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

201-3	Identification of Deposit	
201-3-1	Name of depositary institution	American Type Culture Collection
201-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
201-3-3	Date of deposit	01 July 1998 (01.07.1998)
201-3-4	Accession Number	ATCC 203049
201-4	Additional Indications	NONE
201-5	Designated States for Which Indications are Made	all designated States
201-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
202	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
202-1	page	102
202-2	line	35
202-3	Identification of Deposit	
202-3-1	Name of depositary institution	American Type Culture Collection
202-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
202-3-3	Date of deposit	06 May 1998 (06.05.1998)
202-3-4	Accession Number	ATCC 209846
202-4	Additional Indications	NONE
202-5	Designated States for Which Indications are Made	all designated States
202-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
203	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
203-1	page	102
203-2	line	36
203-3	Identification of Deposit	
203-3-1	Name of depositary institution	American Type Culture Collection
203-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
203-3-3	Date of deposit	12 May 1998 (12.05.1998)
203-3-4	Accession Number	ATCC 209857
203-4	Additional Indications	NONE
203-5	Designated States for Which Indications are Made	all designated States
203-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

204	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
204-1	page	102
204-2	line	37
204-3	Identification of Deposit	
204-3-1	Name of depositary institution	American Type Culture Collection
204-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
204-3-3	Date of deposit	14 May 1998 (14.05.1998)
204-3-4	Accession Number	ATCC 209864
204-4	Additional Indications	NONE
204-5	Designated States for Which Indications are Made	all designated States
204-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
205	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
205-1	page	102
205-2	line	38
205-3	Identification of Deposit	
205-3-1	Name of depositary institution	American Type Culture Collection
205-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
205-3-3	Date of deposit	20 May 1998 (20.05.1998)
205-3-4	Accession Number	ATCC 209880
205-4	Additional Indications	NONE
205-5	Designated States for Which Indications are Made	all designated States
205-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
206	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
206-1	page	102
206-2	line	39
206-3	Identification of Deposit	
206-3-1	Name of depositary institution	American Type Culture Collection
206-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
206-3-3	Date of deposit	14 May 1998 (14.05.1998)
206-3-4	Accession Number	ATCC 209869
206-4	Additional Indications	NONE
206-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

206-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
207	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
207-1	page	102
207-2	line	40
207-3	Identification of Deposit	
207-3-1	Name of depositary institution	American Type Culture Collection
207-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
207-3-3	Date of deposit	09 June 1998 (09.06.1998)
207-3-4	Accession Number	ATCC 209950
207-4	Additional Indications	NONE
207-5	Designated States for Which Indications are Made	all designated States
207-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
208	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
208-1	page	102
208-2	line	41
208-3	Identification of Deposit	
208-3-1	Name of depositary institution	American Type Culture Collection
208-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
208-3-3	Date of deposit	23 June 1998 (23.06.1998)
208-3-4	Accession Number	ATCC 203008
208-4	Additional Indications	NONE
208-5	Designated States for Which Indications are Made	all designated States
208-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
209	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
209-1	page	102
209-2	line	42
209-3	Identification of Deposit	
209-3-1	Name of depositary institution	American Type Culture Collection
209-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
209-3-3	Date of deposit	23 June 1998 (23.06.1998)
209-3-4	Accession Number	ATCC 203014

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

209-4	Additional Indications	NONE
209-5	Designated States for Which Indications are Made	all designated States
209-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
210	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
210-1	page	102
210-2	line	43
210-3	Identification of Deposit	
210-3-1	Name of depositary institution	American Type Culture Collection
210-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
210-3-3	Date of deposit	11 August 1998 (11.08.1998)
210-3-4	Accession Number	ATCC 203110
210-4	Additional Indications	NONE
210-5	Designated States for Which Indications are Made	all designated States
210-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
211	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
211-1	page	102
211-2	line	44
211-3	Identification of Deposit	
211-3-1	Name of depositary institution	American Type Culture Collection
211-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
211-3-3	Date of deposit	23 June 1998 (23.06.1998)
211-3-4	Accession Number	ATCC 203009
211-4	Additional Indications	NONE
211-5	Designated States for Which Indications are Made	all designated States
211-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
212	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
212-1	page	102
212-2	line	45

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

212-3	Identification of Deposit	
212-3-1	Name of depositary institution	American Type Culture Collection
212-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
212-3-3	Date of deposit	09 June 1998 (09.06.1998)
212-3-4	Accession Number	ATCC 209961
212-4	Additional Indications	NONE
212-5	Designated States for Which Indications are Made	all designated States
212-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
213	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
213-1	page	102
213-2	line	46
213-3	Identification of Deposit	
213-3-1	Name of depositary institution	American Type Culture Collection
213-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
213-3-3	Date of deposit	09 June 1998 (09.06.1998)
213-3-4	Accession Number	ATCC 209962
213-4	Additional Indications	NONE
213-5	Designated States for Which Indications are Made	all designated States
213-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
214	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
214-1	page	102
214-2	line	47
214-3	Identification of Deposit	
214-3-1	Name of depositary institution	American Type Culture Collection
214-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
214-3-3	Date of deposit	14 May 1998 (14.05.1998)
214-3-4	Accession Number	ATCC 209866
214-4	Additional Indications	NONE
214-5	Designated States for Which Indications are Made	all designated States
214-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

215	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
215-1	page	102
215-2	line	48
215-3	Identification of Deposit	
215-3-1	Name of depositary institution	American Type Culture Collection
215-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
215-3-3	Date of deposit	25 August 1998 (25.08.1998)
215-3-4	Accession Number	ATCC 203157
215-4	Additional Indications	NONE
215-5	Designated States for Which Indications are Made	all designated States
215-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
216	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
216-1	page	102
216-2	line	49
216-3	Identification of Deposit	
216-3-1	Name of depositary institution	American Type Culture Collection
216-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
216-3-3	Date of deposit	11 August 1998 (11.08.1998)
216-3-4	Accession Number	ATCC 203106
216-4	Additional Indications	NONE
216-5	Designated States for Which Indications are Made	all designated States
216-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
217	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
217-1	page	102
217-2	line	50
217-3	Identification of Deposit	
217-3-1	Name of depositary institution	American Type Culture Collection
217-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
217-3-3	Date of deposit	09 June 1998 (09.06.1998)
217-3-4	Accession Number	ATCC 209945
217-4	Additional Indications	NONE
217-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

217-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
218	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
218-1	page	102
218-2	line	51
218-3	Identification of Deposit	
218-3-1	Name of depositary institution	American Type Culture Collection
218-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
218-3-3	Date of deposit	16 June 1998 (16.06.1998)
218-3-4	Accession Number	ATCC 209989
218-4	Additional Indications	NONE
218-5	Designated States for Which Indications are Made	all designated States
218-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
219	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
219-1	page	102
219-2	line	52
219-3	Identification of Deposit	
219-3-1	Name of depositary institution	American Type Culture Collection
219-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
219-3-3	Date of deposit	11 August 1998 (11.08.1998)
219-3-4	Accession Number	ATCC 203108
219-4	Additional Indications	NONE
219-5	Designated States for Which Indications are Made	all designated States
219-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
220	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
220-1	page	102
220-2	line	53
220-3	Identification of Deposit	
220-3-1	Name of depositary institution	American Type Culture Collection
220-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
220-3-3	Date of deposit	11 August 1998 (11.08.1998)
220-3-4	Accession Number	ATCC 203111

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

220-4	Additional Indications	NONE
220-5	Designated States for Which Indications are Made	all designated States
220-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
221	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
221-1	page	102
221-2	line	54
221-3	Identification of Deposit	
221-3-1	Name of depositary institution	American Type Culture Collection
221-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
221-3-3	Date of deposit	20 October 1998 (20.10.1998)
221-3-4	Accession Number	ATCC 203359
221-4	Additional Indications	NONE
221-5	Designated States for Which Indications are Made	all designated States
221-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
222	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
222-1	page	102
222-2	line	55
222-3	Identification of Deposit	
222-3-1	Name of depositary institution	American Type Culture Collection
222-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
222-3-3	Date of deposit	16 June 1998 (16.06.1998)
222-3-4	Accession Number	ATCC 209988
222-4	Additional Indications	NONE
222-5	Designated States for Which Indications are Made	all designated States
222-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
223	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
223-1	page	103
223-2	line	2

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

223-3	Identification of Deposit	
223-3-1	Name of depositary institution	American Type Culture Collection
223-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
223-3-3	Date of deposit	16 June 1998 (16.06.1998)
223-3-4	Accession Number	ATCC 209978
223-4	Additional Indications	NONE
223-5	Designated States for Which Indications are Made	all designated States
223-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
224	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
224-1	page	103
224-2	line	3
224-3	Identification of Deposit	
224-3-1	Name of depositary institution	American Type Culture Collection
224-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
224-3-3	Date of deposit	04 August 1998 (04.08.1998)
224-3-4	Accession Number	ATCC 203098
224-4	Additional Indications	NONE
224-5	Designated States for Which Indications are Made	all designated States
224-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
225	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
225-1	page	103
225-2	line	4
225-3	Identification of Deposit	
225-3-1	Name of depositary institution	American Type Culture Collection
225-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
225-3-3	Date of deposit	16 June 1998 (16.06.1998)
225-3-4	Accession Number	ATCC 209980
225-4	Additional Indications	NONE
225-5	Designated States for Which Indications are Made	all designated States
225-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

226	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
226-1	page	103
226-2	line	5
226-3	Identification of Deposit	
226-3-1	Name of depositary institution	American Type Culture Collection
226-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
226-3-3	Date of deposit	04 August 1998 (04.08.1998)
226-3-4	Accession Number	ATCC 203091
226-4	Additional Indications	NONE
226-5	Designated States for Which Indications are Made	all designated States
226-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
227	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
227-1	page	103
227-2	line	6
227-3	Identification of Deposit	
227-3-1	Name of depositary institution	American Type Culture Collection
227-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
227-3-3	Date of deposit	04 August 1998 (04.08.1998)
227-3-4	Accession Number	ATCC 203090
227-4	Additional Indications	NONE
227-5	Designated States for Which Indications are Made	all designated States
227-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
228	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
228-1	page	103
228-2	line	7
228-3	Identification of Deposit	
228-3-1	Name of depositary institution	American Type Culture Collection
228-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
228-3-3	Date of deposit	04 August 1998 (04.08.1998)
228-3-4	Accession Number	ATCC 203092
228-4	Additional Indications	NONE
228-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

228-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
229	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
229-1	page	103
229-2	line	8
229-3	Identification of Deposit	
229-3-1	Name of depositary institution	American Type Culture Collection
229-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
229-3-3	Date of deposit	10 November 1998 (10.11.1998)
229-3-4	Accession Number	ATCC 203452
229-4	Additional Indications	NONE
229-5	Designated States for Which Indications are Made	all designated States
229-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
230	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
230-1	page	103
230-2	line	9
230-3	Identification of Deposit	
230-3-1	Name of depositary institution	American Type Culture Collection
230-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
230-3-3	Date of deposit	01 September 1998 (01.09.1998)
230-3-4	Accession Number	ATCC 203173
230-4	Additional Indications	NONE
230-5	Designated States for Which Indications are Made	all designated States
230-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
231	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
231-1	page	103
231-2	line	10
231-3	Identification of Deposit	
231-3-1	Name of depositary institution	American Type Culture Collection
231-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
231-3-3	Date of deposit	17 November 1998 (17.11.1998)
231-3-4	Accession Number	ATCC 203464

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

231-4	Additional Indications	NONE
231-5	Designated States for Which Indications are Made	all designated States
231-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
232	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
232-1	page	103
232-2	line	11
232-3	Identification of Deposit	
232-3-1	Name of depositary institution	American Type Culture Collection
232-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
232-3-3	Date of deposit	18 August 1998 (18.08.1998)
232-3-4	Accession Number	ATCC 203132
232-4	Additional Indications	NONE
232-5	Designated States for Which Indications are Made	all designated States
232-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
233	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
233-1	page	103
233-2	line	12
233-3	Identification of Deposit	
233-3-1	Name of depositary institution	American Type Culture Collection
233-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
233-3-3	Date of deposit	09 September 1998 (09.09.1998)
233-3-4	Accession Number	ATCC 203254
233-4	Additional Indications	NONE
233-5	Designated States for Which Indications are Made	all designated States
233-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
234	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
234-1	page	103
234-2	line	13

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

234-3	Identification of Deposit	
234-3-1	Name of depositary institution	American Type Culture Collection
234-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
234-3-3	Date of deposit	20 October 1998 (20.10.1998)
234-3-4	Accession Number	ATCC 203358
234-4	Additional Indications	NONE
234-5	Designated States for Which Indications are Made	all designated States
234-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
235	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
235-1	page	103
235-2	line	14
235-3	Identification of Deposit	
235-3-1	Name of depositary institution	American Type Culture Collection
235-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
235-3-3	Date of deposit	04 August 1998 (04.08.1998)
235-3-4	Accession Number	ATCC 203093
235-4	Additional Indications	NONE
235-5	Designated States for Which Indications are Made	all designated States
235-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
236	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
236-1	page	103
236-2	line	15
236-3	Identification of Deposit	
236-3-1	Name of depositary institution	American Type Culture Collection
236-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
236-3-3	Date of deposit	03 November 1998 (03.11.1998)
236-3-4	Accession Number	ATCC 203457
236-4	Additional Indications	NONE
236-5	Designated States for Which Indications are Made	all designated States
236-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

237	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
237-1	page	103
237-2	line	16
237-3	Identification of Deposit	
237-3-1	Name of depositary institution	American Type Culture Collection
237-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
237-3-3	Date of deposit	09 September 1998 (09.09.1998)
237-3-4	Accession Number	ATCC 203241
237-4	Additional Indications	NONE
237-5	Designated States for Which Indications are Made	all designated States
237-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
238	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
238-1	page	103
238-2	line	17
238-3	Identification of Deposit	
238-3-1	Name of depositary institution	American Type Culture Collection
238-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
238-3-3	Date of deposit	09 September 1998 (09.09.1998)
238-3-4	Accession Number	ATCC 203249
238-4	Additional Indications	NONE
238-5	Designated States for Which Indications are Made	all designated States
238-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
239	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
239-1	page	103
239-2	line	18
239-3	Identification of Deposit	
239-3-1	Name of depositary institution	American Type Culture Collection
239-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
239-3-3	Date of deposit	09 September 1998 (09.09.1998)
239-3-4	Accession Number	ATCC 203250
239-4	Additional Indications	NONE
239-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

239-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
240	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
240-1	page	103
240-2	line	19
240-3	Identification of Deposit	
240-3-1	Name of depositary institution	American Type Culture Collection
240-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
240-3-3	Date of deposit	18 August 1998 (18.08.1998)
240-3-4	Accession Number	ATCC 203131
240-4	Additional Indications	NONE
240-5	Designated States for Which Indications are Made	all designated States
240-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
241	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
241-1	page	103
241-2	line	20
241-3	Identification of Deposit	
241-3-1	Name of depositary institution	American Type Culture Collection
241-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
241-3-3	Date of deposit	15 September 1998 (15.09.1998)
241-3-4	Accession Number	ATCC 203223
241-4	Additional Indications	NONE
241-5	Designated States for Which Indications are Made	all designated States
241-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
242	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
242-1	page	103
242-2	line	21
242-3	Identification of Deposit	
242-3-1	Name of depositary institution	American Type Culture Collection
242-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
242-3-3	Date of deposit	15 September 1998 (15.09.1998)
242-3-4	Accession Number	ATCC 203233

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

242-4	Additional Indications	NONE
242-5	Designated States for Which Indications are Made	all designated States
242-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
243	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
243-1	page	103
243-2	line	22
243-3	Identification of Deposit	
243-3-1	Name of depositary institution	American Type Culture Collection
243-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
243-3-3	Date of deposit	09 September 1998 (09.09.1998)
243-3-4	Accession Number	ATCC 203252
243-4	Additional Indications	NONE
243-5	Designated States for Which Indications are Made	all designated States
243-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
244	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
244-1	page	103
244-2	line	23
244-3	Identification of Deposit	
244-3-1	Name of depositary institution	American Type Culture Collection
244-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
244-3-3	Date of deposit	17 November 1998 (17.11.1998)
244-3-4	Accession Number	ATCC 203476
244-4	Additional Indications	NONE
244-5	Designated States for Which Indications are Made	all designated States
244-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
245	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
245-1	page	103
245-2	line	24

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

245-3	Identification of Deposit	
245-3-1	Name of depositary institution	American Type Culture Collection
245-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
245-3-3	Date of deposit	04 August 1998 (04.08.1998)
245-3-4	Accession Number	ATCC 203094
245-4	Additional Indications	NONE
245-5	Designated States for Which Indications are Made	all designated States
245-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
246	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
246-1	page	103
246-2	line	25
246-3	Identification of Deposit	
246-3-1	Name of depositary institution	American Type Culture Collection
246-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
246-3-3	Date of deposit	15 September 1998 (15.09.1998)
246-3-4	Accession Number	ATCC 203235
246-4	Additional Indications	NONE
246-5	Designated States for Which Indications are Made	all designated States
246-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
247	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
247-1	page	103
247-2	line	26
247-3	Identification of Deposit	
247-3-1	Name of depositary institution	American Type Culture Collection
247-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
247-3-3	Date of deposit	22 September 1998 (22.09.1998)
247-3-4	Accession Number	ATCC 203267
247-4	Additional Indications	NONE
247-5	Designated States for Which Indications are Made	all designated States
247-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

248	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
248-1	page	103
248-2	line	27
248-3	Identification of Deposit	
248-3-1	Name of depositary institution	American Type Culture Collection
248-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
248-3-3	Date of deposit	22 September 1998 (22.09.1998)
248-3-4	Accession Number	ATCC 203282
248-4	Additional Indications	NONE
248-5	Designated States for Which Indications are Made	all designated States
248-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
249	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
249-1	page	103
249-2	line	28
249-3	Identification of Deposit	
249-3-1	Name of depositary institution	American Type Culture Collection
249-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
249-3-3	Date of deposit	09 February 1999 (09.02.1999)
249-3-4	Accession Number	ATCC 203657
249-4	Additional Indications	NONE
249-5	Designated States for Which Indications are Made	all designated States
249-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
250	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
250-1	page	103
250-2	line	29
250-3	Identification of Deposit	
250-3-1	Name of depositary institution	American Type Culture Collection
250-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
250-3-3	Date of deposit	22 September 1998 (22.09.1998)
250-3-4	Accession Number	ATCC 203276
250-4	Additional Indications	NONE
250-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

250-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
251	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
251-1	page	103
251-2	line	30
251-3	Identification of Deposit	
251-3-1	Name of depositary institution	American Type Culture Collection
251-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
251-3-3	Date of deposit	25 August 1998 (25.08.1998)
251-3-4	Accession Number	ATCC 203160
251-4	Additional Indications	NONE
251-5	Designated States for Which Indications are Made	all designated States
251-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
252	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
252-1	page	103
252-2	line	31
252-3	Identification of Deposit	
252-3-1	Name of depositary institution	American Type Culture Collection
252-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
252-3-3	Date of deposit	18 August 1998 (18.08.1998)
252-3-4	Accession Number	ATCC 203135
252-4	Additional Indications	NONE
252-5	Designated States for Which Indications are Made	all designated States
252-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
253	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
253-1	page	103
253-2	line	32
253-3	Identification of Deposit	
253-3-1	Name of depositary institution	American Type Culture Collection
253-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
253-3-3	Date of deposit	03 November 1998 (03.11.1998)
253-3-4	Accession Number	ATCC 203459

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

253-4	Additional Indications	NONE
253-5	Designated States for Which Indications are Made	all designated States
253-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
254	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
254-1	page	103
254-2	line	33
254-3	Identification of Deposit	
254-3-1	Name of depositary institution	American Type Culture Collection
254-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
254-3-3	Date of deposit	22 September 1998 (22.09.1998)
254-3-4	Accession Number	ATCC 203270
254-4	Additional Indications	NONE
254-5	Designated States for Which Indications are Made	all designated States
254-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
255	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
255-1	page	103
255-2	line	34
255-3	Identification of Deposit	
255-3-1	Name of depositary institution	American Type Culture Collection
255-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
255-3-3	Date of deposit	12 January 1999 (12.01.1999)
255-3-4	Accession Number	ATCC 203573
255-4	Additional Indications	NONE
255-5	Designated States for Which Indications are Made	all designated States
255-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
256	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
256-1	page	103
256-2	line	35

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

256-3	Identification of Deposit	
256-3-1	Name of depositary institution	American Type Culture Collection
256-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
256-3-3	Date of deposit	17 November 1998 (17.11.1998)
256-3-4	Accession Number	ATCC 203477
256-4	Additional Indications	NONE
256-5	Designated States for Which Indications are Made	all designated States
256-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
257	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
257-1	page	103
257-2	line	36
257-3	Identification of Deposit	
257-3-1	Name of depositary institution	American Type Culture Collection
257-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
257-3-3	Date of deposit	06 October 1998 (06.10.1998)
257-3-4	Accession Number	ATCC 203315
257-4	Additional Indications	NONE
257-5	Designated States for Which Indications are Made	all designated States
257-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
258	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
258-1	page	103
258-2	line	37
258-3	Identification of Deposit	
258-3-1	Name of depositary institution	American Type Culture Collection
258-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209United States of America
258-3-3	Date of deposit	06 October 1998 (06.10.1998)
258-3-4	Accession Number	ATCC 203313
258-4	Additional Indications	NONE
258-5	Designated States for Which Indications are Made	all designated States
258-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

259	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
259-1	page	103
259-2	line	38
259-3	Identification of Deposit	
259-3-1	Name of depositary institution	American Type Culture Collection
259-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
259-3-3	Date of deposit	27 October 1998 (27.10.1998)
259-3-4	Accession Number	ATCC 203407
259-4	Additional Indications	NONE
259-5	Designated States for Which Indications are Made	all designated States
259-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
260	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
260-1	page	103
260-2	line	39
260-3	Identification of Deposit	
260-3-1	Name of depositary institution	American Type Culture Collection
260-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
260-3-3	Date of deposit	22 December 1998 (22.12.1998)
260-3-4	Accession Number	ATCC 203553
260-4	Additional Indications	NONE
260-5	Designated States for Which Indications are Made	all designated States
260-6	Separate Furnishing of Indications	NONE
	These indications will be submitted to the International Bureau later	
261	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
261-1	page	103
261-2	line	40
261-3	Identification of Deposit	
261-3-1	Name of depositary institution	American Type Culture Collection
261-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
261-3-3	Date of deposit	22 December 1998 (22.12.1998)
261-3-4	Accession Number	ATCC 203549
261-4	Additional Indications	NONE
261-5	Designated States for Which Indications are Made	all designated States

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

261-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
262	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
262-1	page	103
262-2	line	41
262-3	Identification of Deposit	
262-3-1	Name of depositary institution	American Type Culture Collection
262-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
262-3-3	Date of deposit	22 December 1998 (22.12.1998)
262-3-4	Accession Number	ATCC 203550
262-4	Additional Indications	NONE
262-5	Designated States for Which Indications are Made	all designated States
262-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
263	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
263-1	page	103
263-2	line	42
263-3	Identification of Deposit	
263-3-1	Name of depositary institution	American Type Culture Collection
263-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
263-3-3	Date of deposit	08 June 1999 (08.06.1999)
263-3-4	Accession Number	ATCC PTA-204
263-4	Additional Indications	NONE
263-5	Designated States for Which Indications are Made	all designated States
263-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
264	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
264-1	page	103
264-2	line	43
264-3	Identification of Deposit	
264-3-1	Name of depositary institution	American Type Culture Collection
264-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
264-3-3	Date of deposit	29 October 1998 (29.10.1998)
264-3-4	Accession Number	ATCC 203391

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

264-4	Additional Indications	NONE
264-5	Designated States for Which Indications are Made	all designated States
264-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
265	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
265-1	page	103
265-2	line	44
265-3	Identification of Deposit	
265-3-1	Name of depositary institution	American Type Culture Collection
265-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
265-3-3	Date of deposit	23 March 1999 (23.03.1999)
265-3-4	Accession Number	ATCC 203863
265-4	Additional Indications	NONE
265-5	Designated States for Which Indications are Made	all designated States
265-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
266	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
266-1	page	103
266-2	line	45
266-3	Identification of Deposit	
266-3-1	Name of depositary institution	American Type Culture Collection
266-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
266-3-3	Date of deposit	09 March 1999 (09.03.1999)
266-3-4	Accession Number	ATCC 203834
266-4	Additional Indications	NONE
266-5	Designated States for Which Indications are Made	all designated States
266-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
267	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
267-1	page	103
267-2	line	46

PCT

P3330R1

Original (for SUBMISSION) - printed on 01.12.2000 02:57:35 PM

267-3	Identification of Deposit	
267-3-1	Name of depositary institution	American Type Culture Collection
267-3-2	Address of depositary institution	10801 University Blvd., Manassas, Virginia 20110-2209 United States of America
267-3-3	Date of deposit	20 July 1999 (20.07.1999)
267-3-4	Accession Number	ATCC PTA-382
267-4	Additional Indications	NONE
267-5	Designated States for Which Indications are Made	all designated States
267-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

FOR RECEIVING OFFICE USE ONLY

0-4	This form was received with the international application: (yes or no)	
0-4-1	Authorized officer	

FOR INTERNATIONAL BUREAU USE ONLY

0-5	This form was received by the International Bureau on:	
0-5-1	Authorized officer	

WHAT IS CLAIMED IS:

1. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence that encodes an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16),
5 Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54),
10 Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92),
15 Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254)

235

NO:520), Figure 522 (SEQ ID NO:522), Figure 524 (SEQ ID NO:524), Figure 526 (SEQ ID NO:526), Figure 528 (SEQ ID NO:528), Figure 530 (SEQ ID NO:530), Figure 532 (SEQ ID NO:532), Figure 534 (SEQ ID NO:534), Figure 536 (SEQ ID NO:536), Figure 538 (SEQ ID NO:538), Figure 540 (SEQ ID NO:540), Figure 542 (SEQ ID NO:542), Figure 544 (SEQ ID NO:544), Figure 546 (SEQ ID NO:546), Figure 548 (SEQ ID NO:548) and Figure 550 (SEQ ID NO:550).

5

2. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence selected from the group consisting of the nucleotide sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:5), Figure 7 (SEQ ID NO:7), Figure 9 (SEQ ID NO:9), Figure 11 (SEQ ID NO:11), Figure 13 (SEQ ID NO:13), Figure 15 (SEQ ID NO:15), Figure 17 (SEQ ID NO:17), Figure 19 (SEQ ID NO:19), Figure 21 (SEQ ID NO:21), Figure 23 (SEQ ID NO:23), Figure 25 (SEQ ID NO:25), Figure 27 (SEQ ID NO:27), Figure 29 (SEQ ID NO:29), Figure 31 (SEQ ID NO:31), Figure 33 (SEQ ID NO:33), Figure 35 (SEQ ID NO:35), Figure 37 (SEQ ID NO:37), Figure 39 (SEQ ID NO:39), Figure 41 (SEQ ID NO:41), Figure 43 (SEQ ID NO:43), Figure 45 (SEQ ID NO:45), Figure 47 (SEQ ID NO:47), Figure 49 (SEQ ID NO:49), Figure 51 (SEQ ID NO:51), Figure 53 (SEQ ID NO:53), Figure 55 (SEQ ID NO:55), Figure 57 (SEQ ID NO:57), Figure 59 (SEQ ID NO:59), Figure 61 (SEQ ID NO:61), Figure 63 (SEQ ID NO:63), Figure 65 (SEQ ID NO:65), Figure 67 (SEQ ID NO:67), Figure 69 (SEQ ID NO:69), Figure 71 (SEQ ID NO:71), Figure 73 (SEQ ID NO:73), Figure 75 (SEQ ID NO:75), Figure 77 (SEQ ID NO:77), Figure 79 (SEQ ID NO:79), Figure 81 (SEQ ID NO:81), Figure 83 (SEQ ID NO:83), Figure 85 (SEQ ID NO:85), Figure 87 (SEQ ID NO:87), Figure 89 (SEQ ID NO:89), Figure 91 (SEQ ID NO:91), Figure 93 (SEQ ID NO:93), Figure 95 (SEQ ID NO:95), Figure 97 (SEQ ID NO:97), Figure 99 (SEQ ID NO:99), Figure 101 (SEQ ID NO:101), Figure 103 (SEQ ID NO:103), Figure 105 (SEQ ID NO:105), Figure 107 (SEQ ID NO:107), Figure 109 (SEQ ID NO:109), Figure 111 (SEQ ID NO:111), Figure 113 (SEQ ID NO:113), Figure 115 (SEQ ID NO:115), Figure 117 (SEQ ID NO:117), Figure 119 (SEQ ID NO:119), Figure 121 (SEQ ID NO:121), Figure 123 (SEQ ID NO:123), Figure 125 (SEQ ID NO:125), Figure 127 (SEQ ID NO:127), Figure 129 (SEQ ID NO:129), Figure 131 (SEQ ID NO:131), Figure 133 (SEQ ID NO:133), Figure 135 (SEQ ID NO:135), Figure 137 (SEQ ID NO:137), Figure 139 (SEQ ID NO:139), Figure 141 (SEQ ID NO:141), Figure 143 (SEQ ID NO:143), Figure 145 (SEQ ID NO:145), Figure 147 (SEQ ID NO:147), Figure 149 (SEQ ID NO:149), Figure 151 (SEQ ID NO:151), Figure 153 (SEQ ID NO:153), Figure 155 (SEQ ID NO:155), Figure 157 (SEQ ID NO:157), Figure 159 (SEQ ID NO:159), Figure 161 (SEQ ID NO:161), Figure 163 (SEQ ID NO:163), Figure 165 (SEQ ID NO:165), Figure 167 (SEQ ID NO:167), Figure 169 (SEQ ID NO:169), Figure 171 (SEQ ID NO:171), Figure 173 (SEQ ID NO:173), Figure 175 (SEQ ID NO:175), Figure 177 (SEQ ID NO:177), Figure 179 (SEQ ID NO:179), Figure 181 (SEQ ID NO:181), Figure 183 (SEQ ID NO:183), Figure 185 (SEQ ID NO:185), Figure 187 (SEQ ID NO:187), Figure 189 (SEQ ID NO:189), Figure 191 (SEQ ID NO:191), Figure 193 (SEQ ID NO:193), Figure 195 (SEQ ID NO:195), Figure 197 (SEQ ID NO:197), Figure 199 (SEQ ID NO:199), Figure 201 (SEQ ID NO:201), Figure 203 (SEQ ID NO:203), Figure 205 (SEQ ID NO:205), Figure 207 (SEQ ID NO:207), Figure 209 (SEQ ID NO:209), Figure 211 (SEQ ID NO:211), Figure 213 (SEQ ID NO:213), Figure 215 (SEQ ID NO:215), Figure 217 (SEQ ID NO:217), Figure 219 (SEQ ID NO:219), Figure 221 (SEQ ID

237

NO:487), Figure 489 (SEQ ID NO:489), Figure 491 (SEQ ID NO:491), Figure 493 (SEQ ID NO:493), Figure 495 (SEQ ID NO:495), Figure 497 (SEQ ID NO:497), Figure 499 (SEQ ID NO:499), Figure 501 (SEQ ID NO:501), Figure 503 (SEQ ID NO:503), Figure 505 (SEQ ID NO:505), Figure 507 (SEQ ID NO:507), Figure 509 (SEQ ID NO:509), Figure 511 (SEQ ID NO:511), Figure 513 (SEQ ID NO:513), Figure 515 (SEQ ID NO:515), Figure 517 (SEQ ID NO:517), Figure 519 (SEQ ID NO:519), Figure 521 (SEQ ID NO:521), Figure 523 (SEQ ID NO:523), Figure 525 (SEQ ID NO:525), Figure 527 (SEQ ID NO:527), Figure 529 (SEQ ID NO:529), Figure 531 (SEQ ID NO:531), Figure 533 (SEQ ID NO:533), Figure 535 (SEQ ID NO:535), Figure 537 (SEQ ID NO:537), Figure 539 (SEQ ID NO:539), Figure 541 (SEQ ID NO:541), Figure 543 (SEQ ID NO:543), Figure 545 (SEQ ID NO:545), Figure 547 (SEQ ID NO:547) and Figure 549 (SEQ ID NO:549).

3. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence selected from the group consisting of the full-length coding sequence of the nucleotide sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:5), Figure 7 (SEQ ID NO:7), Figure 9 (SEQ ID NO:9), Figure 11 (SEQ ID NO:11), Figure 13 (SEQ ID NO:13), Figure 15 (SEQ ID NO:15), Figure 17 (SEQ ID NO:17), Figure 19 (SEQ ID NO:19), Figure 21 (SEQ ID NO:21), Figure 23 (SEQ ID NO:23), Figure 25 (SEQ ID NO:25), Figure 27 (SEQ ID NO:27), Figure 29 (SEQ ID NO:29), Figure 31 (SEQ ID NO:31), Figure 33 (SEQ ID NO:33), Figure 35 (SEQ ID NO:35), Figure 37 (SEQ ID NO:37), Figure 39 (SEQ ID NO:39), Figure 41 (SEQ ID NO:41), Figure 43 (SEQ ID NO:43), Figure 45 (SEQ ID NO:45), Figure 47 (SEQ ID NO:47), Figure 49 (SEQ ID NO:49), Figure 51 (SEQ ID NO:51), Figure 53 (SEQ ID NO:53), Figure 55 (SEQ ID NO:55), Figure 57 (SEQ ID NO:57), Figure 59 (SEQ ID NO:59), Figure 61 (SEQ ID NO:61), Figure 63 (SEQ ID NO:63), Figure 65 (SEQ ID NO:65), Figure 67 (SEQ ID NO:67), Figure 69 (SEQ ID NO:69), Figure 71 (SEQ ID NO:71), Figure 73 (SEQ ID NO:73), Figure 75 (SEQ ID NO:75), Figure 77 (SEQ ID NO:77), Figure 79 (SEQ ID NO:79), Figure 81 (SEQ ID NO:81), Figure 83 (SEQ ID NO:83), Figure 85 (SEQ ID NO:85), Figure 87 (SEQ ID NO:87), Figure 89 (SEQ ID NO:89), Figure 91 (SEQ ID NO:91), Figure 93 (SEQ ID NO:93), Figure 95 (SEQ ID NO:95), Figure 97 (SEQ ID NO:97), Figure 99 (SEQ ID NO:99), Figure 101 (SEQ ID NO:101), Figure 103 (SEQ ID NO:103), Figure 105 (SEQ ID NO:105), Figure 107 (SEQ ID NO:107), Figure 109 (SEQ ID NO:109), Figure 111 (SEQ ID NO:111), Figure 113 (SEQ ID NO:113), Figure 115 (SEQ ID NO:115), Figure 117 (SEQ ID NO:117), Figure 119 (SEQ ID NO:119), Figure 121 (SEQ ID NO:121), Figure 123 (SEQ ID NO:123), Figure 125 (SEQ ID NO:125), Figure 127 (SEQ ID NO:127), Figure 129 (SEQ ID NO:129), Figure 131 (SEQ ID NO:131), Figure 133 (SEQ ID NO:133), Figure 135 (SEQ ID NO:135), Figure 137 (SEQ ID NO:137), Figure 139 (SEQ ID NO:139), Figure 141 (SEQ ID NO:141), Figure 143 (SEQ ID NO:143), Figure 145 (SEQ ID NO:145), Figure 147 (SEQ ID NO:147), Figure 149 (SEQ ID NO:149), Figure 151 (SEQ ID NO:151), Figure 153 (SEQ ID NO:153), Figure 155 (SEQ ID NO:155), Figure 157 (SEQ ID NO:157), Figure 159 (SEQ ID NO:159), Figure 161 (SEQ ID NO:161), Figure 163 (SEQ ID NO:163), Figure 165 (SEQ ID NO:165), Figure 167 (SEQ ID NO:167), Figure 169 (SEQ ID NO:169), Figure 171 (SEQ ID NO:171), Figure 173 (SEQ ID NO:173), Figure 175 (SEQ ID NO:175), Figure 177 (SEQ ID NO:177), Figure 179 (SEQ ID NO:179), Figure 181 (SEQ ID NO:181), Figure 183 (SEQ ID NO:183), Figure 185 (SEQ ID NO:185), Figure 187 (SEQ ID NO:187), Figure 189 (SEQ ID NO:189), Figure 191 (SEQ

ID NO:191), Figure 193 (SEQ ID NO:193), Figure 195 (SEQ ID NO:195), Figure 197 (SEQ ID NO:197),
Figure 199 (SEQ ID NO:199), Figure 201 (SEQ ID NO:201), Figure 203 (SEQ ID NO:203), Figure 205 (SEQ
ID NO:205), Figure 207 (SEQ ID NO:207), Figure 209 (SEQ ID NO:209), Figure 211 (SEQ ID NO:211),
Figure 213 (SEQ ID NO:213), Figure 215 (SEQ ID NO:215), Figure 217 (SEQ ID NO:217), Figure 219 (SEQ
ID NO:219), Figure 221 (SEQ ID NO:221), Figure 223 (SEQ ID NO:223), Figure 225 (SEQ ID NO:225),
5 Figure 227 (SEQ ID NO:227), Figure 229 (SEQ ID NO:229), Figure 231 (SEQ ID NO:231), Figure 233 (SEQ
ID NO:233), Figure 235 (SEQ ID NO:235), Figure 237 (SEQ ID NO:237), Figure 239 (SEQ ID NO:239),
Figure 241 (SEQ ID NO:241), Figure 243 (SEQ ID NO:243), Figure 245 (SEQ ID NO:245), Figure 247 (SEQ
ID NO:247), Figure 249 (SEQ ID NO:249), Figure 251 (SEQ ID NO:251), Figure 253 (SEQ ID NO:253),
Figure 255 (SEQ ID NO:255), Figure 257 (SEQ ID NO:257), Figure 259 (SEQ ID NO:259), Figure 261 (SEQ
10 ID NO:261), Figure 263 (SEQ ID NO:263), Figure 265 (SEQ ID NO:265), Figure 267 (SEQ ID NO:267),
Figure 269 (SEQ ID NO:269), Figure 271 (SEQ ID NO:271), Figure 273 (SEQ ID NO:273), Figure 275 (SEQ
ID NO:275), Figure 277 (SEQ ID NO:277), Figure 279 (SEQ ID NO:279), Figure 281 (SEQ ID NO:281),
Figure 283 (SEQ ID NO:283), Figure 285 (SEQ ID NO:285), Figure 287 (SEQ ID NO:287), Figure 289 (SEQ
ID NO:289), Figure 291 (SEQ ID NO:291), Figure 293 (SEQ ID NO:293), Figure 295 (SEQ ID NO:295),
15 Figure 297 (SEQ ID NO:297), Figure 299 (SEQ ID NO:299), Figure 301 (SEQ ID NO:301), Figure 303 (SEQ
ID NO:303), Figure 305 (SEQ ID NO:305), Figure 307 (SEQ ID NO:307), Figure 309 (SEQ ID NO:309),
Figure 311 (SEQ ID NO:311), Figure 313 (SEQ ID NO:313), Figure 315 (SEQ ID NO:315), Figure 317 (SEQ
ID NO:317), Figure 319 (SEQ ID NO:319), Figure 321 (SEQ ID NO:321), Figure 323 (SEQ ID NO:323),
Figure 325 (SEQ ID NO:325), Figure 327 (SEQ ID NO:327), Figure 329 (SEQ ID NO:329), Figure 331 (SEQ
20 ID NO:331), Figure 333 (SEQ ID NO:333), Figure 335 (SEQ ID NO:335), Figure 337 (SEQ ID NO:337),
Figure 339 (SEQ ID NO:339), Figure 341 (SEQ ID NO:341), Figure 343 (SEQ ID NO:343), Figure 345 (SEQ
ID NO:345), Figure 347 (SEQ ID NO:347), Figure 349 (SEQ ID NO:349), Figure 351 (SEQ ID NO:351),
Figure 353 (SEQ ID NO:353), Figure 355 (SEQ ID NO:355), Figure 357 (SEQ ID NO:357), Figure 359 (SEQ
ID NO:359), Figure 361 (SEQ ID NO:361), Figure 363 (SEQ ID NO:363), Figure 365 (SEQ ID NO:365),
25 Figure 367 (SEQ ID NO:367), Figure 369 (SEQ ID NO:369), Figure 371 (SEQ ID NO:371), Figure 373 (SEQ
ID NO:373), Figure 375 (SEQ ID NO:375), Figure 377 (SEQ ID NO:377), Figure 379 (SEQ ID NO:379),
Figure 381 (SEQ ID NO:381), Figure 383 (SEQ ID NO:383), Figure 385 (SEQ ID NO:385), Figure 387 (SEQ
ID NO:387), Figure 389 (SEQ ID NO:389), Figure 391 (SEQ ID NO:391), Figure 393 (SEQ ID NO:393),
Figure 395 (SEQ ID NO:395), Figure 397 (SEQ ID NO:397), Figure 399 (SEQ ID NO:399), Figure 401 (SEQ
30 ID NO:401), Figure 403 (SEQ ID NO:403), Figure 405 (SEQ ID NO:405), Figure 407 (SEQ ID NO:407),
Figure 409 (SEQ ID NO:409), Figure 411 (SEQ ID NO:411), Figure 413 (SEQ ID NO:413), Figure 415 (SEQ
ID NO:415), Figure 417 (SEQ ID NO:417), Figure 419 (SEQ ID NO:419), Figure 421 (SEQ ID NO:421),
Figure 423 (SEQ ID NO:423), Figure 425 (SEQ ID NO:425), Figure 427 (SEQ ID NO:427), Figure 429 (SEQ
ID NO:429), Figure 431 (SEQ ID NO:431), Figure 433 (SEQ ID NO:433), Figure 435 (SEQ ID NO:435),
35 Figure 437 (SEQ ID NO:437), Figure 439 (SEQ ID NO:439), Figure 441 (SEQ ID NO:441), Figure 443 (SEQ
ID NO:443), Figure 445 (SEQ ID NO:445), Figure 447 (SEQ ID NO:447), Figure 449 (SEQ ID NO:449),
Figure 451 (SEQ ID NO:451), Figure 453 (SEQ ID NO:453), Figure 455 (SEQ ID NO:455), Figure 457 (SEQ

ID NO:457), Figure 459 (SEQ ID NO:459), Figure 461 (SEQ ID NO:461), Figure 463 (SEQ ID NO:463), Figure 465 (SEQ ID NO:465), Figure 467 (SEQ ID NO:467), Figure 469 (SEQ ID NO:469), Figure 471 (SEQ ID NO:471), Figure 473 (SEQ ID NO:473), Figure 475 (SEQ ID NO:475), Figure 477 (SEQ ID NO:477), Figure 479 (SEQ ID NO:479), Figure 481 (SEQ ID NO:481), Figure 483 (SEQ ID NO:483), Figure 485 (SEQ ID NO:485), Figure 487 (SEQ ID NO:487), Figure 489 (SEQ ID NO:489), Figure 491 (SEQ ID NO:491),
5 Figure 493 (SEQ ID NO:493), Figure 495 (SEQ ID NO:495), Figure 497 (SEQ ID NO:497), Figure 499 (SEQ ID NO:499), Figure 501 (SEQ ID NO:501), Figure 503 (SEQ ID NO:503), Figure 505 (SEQ ID NO:505), Figure 507 (SEQ ID NO:507), Figure 509 (SEQ ID NO:509), Figure 511 (SEQ ID NO:511), Figure 513 (SEQ ID NO:513), Figure 515 (SEQ ID NO:515), Figure 517 (SEQ ID NO:517), Figure 519 (SEQ ID NO:519), Figure 521 (SEQ ID NO:521), Figure 523 (SEQ ID NO:523), Figure 525 (SEQ ID NO:525), Figure 527 (SEQ ID NO:527), Figure 529 (SEQ ID NO:529), Figure 531 (SEQ ID NO:531), Figure 533 (SEQ ID NO:533),
10 Figure 535 (SEQ ID NO:535), Figure 537 (SEQ ID NO:537), Figure 539 (SEQ ID NO:539), Figure 541 (SEQ ID NO:541), Figure 543 (SEQ ID NO:543), Figure 545 (SEQ ID NO:545), Figure 547 (SEQ ID NO:547) and Figure 549 (SEQ ID NO:549).

15 4. Isolated nucleic acid having at least 80% nucleic acid sequence identity to the full-length coding sequence of the DNA deposited under any ATCC accession number shown in Table 7.

5. A vector comprising the nucleic acid of Claim 1.

20 6. The vector of Claim 5 operably linked to control sequences recognized by a host cell transformed with the vector.

7. A host cell comprising the vector of Claim 5.

25 8. The host cell of Claim 7, wherein said cell is a CHO cell.

9. The host cell of Claim 7, wherein said cell is an *E. coli*.

10. The host cell of Claim 7, wherein said cell is a yeast cell.

30 11. A process for producing a PRO polypeptides comprising culturing the host cell of Claim 7 under conditions suitable for expression of said PRO polypeptide and recovering said PRO polypeptide from the cell culture.

35 12. An isolated polypeptide having at least 80% amino acid sequence identity to an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10),

Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48),
5 Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86),
10 Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure

284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372), Figure 374 (SEQ ID NO:374), Figure 376 (SEQ ID NO:376), Figure 378 (SEQ ID NO:378), Figure 380 (SEQ ID NO:380), Figure 382 (SEQ ID NO:382), Figure 384 (SEQ ID NO:384), Figure 386 (SEQ ID NO:386), Figure 388 (SEQ ID NO:388), Figure 390 (SEQ ID NO:390), Figure 392 (SEQ ID NO:392), Figure 394 (SEQ ID NO:394), Figure 396 (SEQ ID NO:396), Figure 398 (SEQ ID NO:398), Figure 400 (SEQ ID NO:400), Figure 402 (SEQ ID NO:402), Figure 404 (SEQ ID NO:404), Figure 406 (SEQ ID NO:406), Figure 408 (SEQ ID NO:408), Figure 410 (SEQ ID NO:410), Figure 412 (SEQ ID NO:412), Figure 414 (SEQ ID NO:414), Figure 416 (SEQ ID NO:416), Figure 418 (SEQ ID NO:418), Figure 420 (SEQ ID NO:420), Figure 422 (SEQ ID NO:422), Figure 424 (SEQ ID NO:424), Figure 426 (SEQ ID NO:426), Figure 428 (SEQ ID NO:428), Figure 430 (SEQ ID NO:430), Figure 432 (SEQ ID NO:432), Figure 434 (SEQ ID NO:434), Figure 436 (SEQ ID NO:436), Figure 438 (SEQ ID NO:438), Figure 440 (SEQ ID NO:440), Figure 442 (SEQ ID NO:442), Figure 444 (SEQ ID NO:444), Figure 446 (SEQ ID NO:446), Figure 448 (SEQ ID NO:448), Figure 450 (SEQ ID NO:450), Figure 452 (SEQ ID NO:452), Figure 454 (SEQ ID NO:454), Figure 456 (SEQ ID NO:456), Figure 458 (SEQ ID NO:458), Figure 460 (SEQ ID NO:460), Figure 462 (SEQ ID NO:462), Figure 464 (SEQ ID NO:464), Figure 466 (SEQ ID NO:466), Figure 468 (SEQ ID NO:468), Figure 470 (SEQ ID NO:470), Figure 472 (SEQ ID NO:472), Figure 474 (SEQ ID NO:474), Figure 476 (SEQ ID NO:476), Figure 478 (SEQ ID NO:478), Figure 480 (SEQ ID NO:480), Figure 482 (SEQ ID NO:482), Figure 484 (SEQ ID NO:484), Figure 486 (SEQ ID NO:486), Figure 488 (SEQ ID NO:488), Figure 490 (SEQ ID NO:490), Figure 492 (SEQ ID NO:492), Figure 494 (SEQ ID NO:494), Figure 496 (SEQ ID NO:496), Figure 498 (SEQ ID NO:498), Figure 500 (SEQ ID NO:500), Figure 502 (SEQ ID NO:502), Figure 504 (SEQ ID NO:504), Figure 506 (SEQ ID NO:506), Figure 508 (SEQ ID NO:508), Figure 510 (SEQ ID NO:510), Figure 512 (SEQ ID NO:512), Figure 514 (SEQ ID NO:514), Figure 516 (SEQ ID NO:516), Figure 518 (SEQ ID NO:518), Figure 520 (SEQ ID NO:520), Figure 522 (SEQ ID NO:522), Figure 524 (SEQ ID NO:524), Figure 526 (SEQ ID NO:526), Figure 528 (SEQ ID NO:528), Figure 530 (SEQ ID NO:530), Figure 532 (SEQ ID NO:532), Figure 534 (SEQ ID NO:534), Figure 536 (SEQ ID NO:536), Figure 538 (SEQ ID NO:538), Figure 540 (SEQ ID NO:540), Figure 542 (SEQ ID NO:542), Figure 544 (SEQ ID NO:544), Figure 546 (SEQ ID NO:546), Figure 548 (SEQ ID NO:548) and

Figure 550 (SEQ ID NO:550).

13. An isolated polypeptide having at least 80% amino acid sequence identity to an amino acid sequence encoded by the full-length coding sequence of the DNA deposited under any ATCC accession number shown in Table 7.

14. A chimeric molecule comprising a polypeptide according to Claim 12 fused to a heterologous amino acid sequence.

15. The chimeric molecule of Claim 14, wherein said heterologous amino acid sequence is an epitope tag sequence.

16. The chimeric molecule of Claim 14, wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

17. An antibody which specifically binds to a polypeptide according to Claim 12.

18. The antibody of Claim 17, wherein said antibody is a monoclonal antibody, a humanized antibody or a single-chain antibody.

19. Isolated nucleic acid having at least 80% nucleic acid sequence identity to:

(a) a nucleotide sequence encoding the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ

ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158),
5 Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200),
10 Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228),
15 Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270),
20 Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298),
25 Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340),
30 Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368),
35 Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372), Figure 374 (SEQ ID NO:374), Figure 376 (SEQ ID NO:376), Figure 378 (SEQ ID NO:378), Figure 380 (SEQ ID NO:380), Figure 382 (SEQ ID NO:382), Figure 384 (SEQ ID NO:384), Figure 386 (SEQ ID NO:386), Figure 388 (SEQ ID NO:388), Figure 390 (SEQ

ID NO:390), Figure 392 (SEQ ID NO:392), Figure 394 (SEQ ID NO:394), Figure 396 (SEQ ID NO:396), Figure 398 (SEQ ID NO:398), Figure 400 (SEQ ID NO:400), Figure 402 (SEQ ID NO:402), Figure 404 (SEQ ID NO:404), Figure 406 (SEQ ID NO:406), Figure 408 (SEQ ID NO:408), Figure 410 (SEQ ID NO:410), Figure 412 (SEQ ID NO:412), Figure 414 (SEQ ID NO:414), Figure 416 (SEQ ID NO:416), Figure 418 (SEQ ID NO:418), Figure 420 (SEQ ID NO:420), Figure 422 (SEQ ID NO:422), Figure 424 (SEQ ID NO:424),
 5 Figure 426 (SEQ ID NO:426), Figure 428 (SEQ ID NO:428), Figure 430 (SEQ ID NO:430), Figure 432 (SEQ ID NO:432), Figure 434 (SEQ ID NO:434), Figure 436 (SEQ ID NO:436), Figure 438 (SEQ ID NO:438), Figure 440 (SEQ ID NO:440), Figure 442 (SEQ ID NO:442), Figure 444 (SEQ ID NO:444), Figure 446 (SEQ ID NO:446), Figure 448 (SEQ ID NO:448), Figure 450 (SEQ ID NO:450), Figure 452 (SEQ ID NO:452), Figure 454 (SEQ ID NO:454), Figure 456 (SEQ ID NO:456), Figure 458 (SEQ ID NO:458), Figure 460 (SEQ ID NO:460), Figure 462 (SEQ ID NO:462), Figure 464 (SEQ ID NO:464), Figure 466 (SEQ ID NO:466), Figure 468 (SEQ ID NO:468), Figure 470 (SEQ ID NO:470), Figure 472 (SEQ ID NO:472), Figure 474 (SEQ ID NO:474), Figure 476 (SEQ ID NO:476), Figure 478 (SEQ ID NO:478), Figure 480 (SEQ ID NO:480), Figure 482 (SEQ ID NO:482), Figure 484 (SEQ ID NO:484), Figure 486 (SEQ ID NO:486), Figure 488 (SEQ ID NO:488), Figure 490 (SEQ ID NO:490), Figure 492 (SEQ ID NO:492), Figure 494 (SEQ ID NO:494),
 15 Figure 496 (SEQ ID NO:496), Figure 498 (SEQ ID NO:498), Figure 500 (SEQ ID NO:500), Figure 502 (SEQ ID NO:502), Figure 504 (SEQ ID NO:504), Figure 506 (SEQ ID NO:506), Figure 508 (SEQ ID NO:508), Figure 510 (SEQ ID NO:510), Figure 512 (SEQ ID NO:512), Figure 514 (SEQ ID NO:514), Figure 516 (SEQ ID NO:516), Figure 518 (SEQ ID NO:518), Figure 520 (SEQ ID NO:520), Figure 522 (SEQ ID NO:522), Figure 524 (SEQ ID NO:524), Figure 526 (SEQ ID NO:526), Figure 528 (SEQ ID NO:528), Figure 530 (SEQ ID NO:530), Figure 532 (SEQ ID NO:532), Figure 534 (SEQ ID NO:534), Figure 536 (SEQ ID NO:536), Figure 538 (SEQ ID NO:538), Figure 540 (SEQ ID NO:540), Figure 542 (SEQ ID NO:542), Figure 544 (SEQ ID NO:544), Figure 546 (SEQ ID NO:546), Figure 548 (SEQ ID NO:548) or Figure 550 (SEQ ID NO:550), lacking its associated signal peptide;

(b) a nucleotide sequence encoding an extracellular domain of the polypeptide shown in Figure 2
 25 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100),
 35

246

Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372), Figure 374 (SEQ ID NO:374), Figure 376 (SEQ ID NO:376), Figure 378 (SEQ ID NO:378), Figure 380 (SEQ ID NO:380), Figure 382 (SEQ ID NO:382), Figure 384 (SEQ ID NO:384), Figure 386 (SEQ ID NO:386), Figure 388 (SEQ ID NO:388), Figure 390 (SEQ ID NO:390), Figure 392 (SEQ ID NO:392), Figure 394 (SEQ ID NO:394), Figure 396 (SEQ ID NO:396), Figure 398 (SEQ ID NO:398), Figure 400 (SEQ ID NO:400), Figure 402 (SEQ ID NO:402), Figure 404 (SEQ ID NO:404), Figure 406 (SEQ ID NO:406), Figure 408 (SEQ ID NO:408), Figure 410 (SEQ ID NO:410), Figure 412 (SEQ ID NO:412), Figure 414 (SEQ ID NO:414), Figure 416 (SEQ ID NO:416), Figure 418 (SEQ ID NO:418), Figure 420 (SEQ ID NO:420), Figure 422 (SEQ ID NO:422), Figure 424 (SEQ ID NO:424), Figure 426 (SEQ ID NO:426), Figure 428 (SEQ ID NO:428), Figure 430 (SEQ ID NO:430), Figure 432 (SEQ ID NO:432), Figure 434 (SEQ ID NO:434), Figure 436 (SEQ ID NO:436), Figure 438 (SEQ ID NO:438), Figure 440 (SEQ ID NO:440), Figure 442 (SEQ ID NO:442), Figure 444 (SEQ ID NO:444), Figure 446 (SEQ ID NO:446), Figure 448 (SEQ ID NO:448), Figure 450 (SEQ ID NO:450), Figure 452 (SEQ ID NO:452), Figure 454 (SEQ ID NO:454), Figure 456 (SEQ ID NO:456), Figure 458 (SEQ ID NO:458), Figure 460 (SEQ ID NO:460), Figure 462 (SEQ ID NO:462), Figure 464 (SEQ ID NO:464), Figure 466 (SEQ ID NO:466), Figure 468 (SEQ ID NO:468), Figure 470 (SEQ ID NO:470), Figure 472 (SEQ ID NO:472), Figure 474 (SEQ ID NO:474), Figure 476 (SEQ ID NO:476), Figure 478 (SEQ ID NO:478), Figure 480 (SEQ ID NO:480), Figure 482 (SEQ ID NO:482), Figure 484 (SEQ ID NO:484), Figure 486 (SEQ ID NO:486), Figure 488 (SEQ ID NO:488), Figure 490 (SEQ ID NO:490), Figure 492 (SEQ ID NO:492), Figure 494 (SEQ ID NO:494), Figure 496 (SEQ ID NO:496), Figure 498 (SEQ ID NO:498), Figure 500 (SEQ ID NO:500), Figure 502 (SEQ ID NO:502), Figure 504 (SEQ ID NO:504), Figure 506 (SEQ ID NO:506), Figure 508 (SEQ ID NO:508), Figure 510 (SEQ ID NO:510), Figure 512 (SEQ ID NO:512), Figure 514 (SEQ ID NO:514), Figure 516 (SEQ ID NO:516), Figure 518 (SEQ ID NO:518), Figure 520 (SEQ ID NO:520), Figure 522 (SEQ ID NO:522), Figure 524 (SEQ ID NO:524), Figure 526 (SEQ ID NO:526), Figure 528 (SEQ ID NO:528), Figure 530 (SEQ ID NO:530), Figure 532 (SEQ ID NO:532), Figure 534 (SEQ ID NO:534), Figure 536 (SEQ ID NO:536), Figure 538 (SEQ ID NO:538), Figure 540 (SEQ ID NO:540), Figure 542 (SEQ ID NO:542), Figure 544 (SEQ ID NO:544), Figure 546 (SEQ ID NO:546), Figure 548 (SEQ ID NO:548) or Figure 550 (SEQ ID NO:550), with its associated signal peptide; or

(c) a nucleotide sequence encoding an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID

NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114),
5 Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184),
15 Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254),
25 Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324),
30 Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ

ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372), Figure 374 (SEQ ID NO:374), Figure 376 (SEQ ID NO:376), Figure 378 (SEQ ID NO:378), Figure 380 (SEQ ID NO:380), Figure 382 (SEQ ID NO:382), Figure 384 (SEQ ID NO:384), Figure 386 (SEQ ID NO:386), Figure 388 (SEQ ID NO:388), Figure 390 (SEQ ID NO:390), Figure 392 (SEQ ID NO:392), Figure 394 (SEQ ID NO:394), Figure 396 (SEQ ID NO:396), Figure 398 (SEQ ID NO:398), Figure 400 (SEQ ID NO:400), Figure 402 (SEQ ID NO:402), Figure 404 (SEQ ID NO:404), Figure 406 (SEQ ID NO:406), Figure 408 (SEQ ID NO:408), Figure 410 (SEQ ID NO:410), Figure 412 (SEQ ID NO:412), Figure 414 (SEQ ID NO:414), Figure 416 (SEQ ID NO:416), Figure 418 (SEQ ID NO:418), Figure 420 (SEQ ID NO:420), Figure 422 (SEQ ID NO:422), Figure 424 (SEQ ID NO:424), Figure 426 (SEQ ID NO:426), Figure 428 (SEQ ID NO:428), Figure 430 (SEQ ID NO:430), Figure 432 (SEQ ID NO:432), Figure 434 (SEQ ID NO:434), Figure 436 (SEQ ID NO:436), Figure 438 (SEQ ID NO:438), Figure 440 (SEQ ID NO:440), Figure 442 (SEQ ID NO:442), Figure 444 (SEQ ID NO:444), Figure 446 (SEQ ID NO:446), Figure 448 (SEQ ID NO:448), Figure 450 (SEQ ID NO:450), Figure 452 (SEQ ID NO:452), Figure 454 (SEQ ID NO:454), Figure 456 (SEQ ID NO:456), Figure 458 (SEQ ID NO:458), Figure 460 (SEQ ID NO:460), Figure 462 (SEQ ID NO:462), Figure 464 (SEQ ID NO:464), Figure 466 (SEQ ID NO:466), Figure 468 (SEQ ID NO:468), Figure 470 (SEQ ID NO:470), Figure 472 (SEQ ID NO:472), Figure 474 (SEQ ID NO:474), Figure 476 (SEQ ID NO:476), Figure 478 (SEQ ID NO:478), Figure 480 (SEQ ID NO:480), Figure 482 (SEQ ID NO:482), Figure 484 (SEQ ID NO:484), Figure 486 (SEQ ID NO:486), Figure 488 (SEQ ID NO:488), Figure 490 (SEQ ID NO:490), Figure 492 (SEQ ID NO:492), Figure 494 (SEQ ID NO:494), Figure 496 (SEQ ID NO:496), Figure 498 (SEQ ID NO:498), Figure 500 (SEQ ID NO:500), Figure 502 (SEQ ID NO:502), Figure 504 (SEQ ID NO:504), Figure 506 (SEQ ID NO:506), Figure 508 (SEQ ID NO:508), Figure 510 (SEQ ID NO:510), Figure 512 (SEQ ID NO:512), Figure 514 (SEQ ID NO:514), Figure 516 (SEQ ID NO:516), Figure 518 (SEQ ID NO:518), Figure 520 (SEQ ID NO:520), Figure 522 (SEQ ID NO:522), Figure 524 (SEQ ID NO:524), Figure 526 (SEQ ID NO:526), Figure 528 (SEQ ID NO:528), Figure 530 (SEQ ID NO:530), Figure 532 (SEQ ID NO:532), Figure 534 (SEQ ID NO:534), Figure 536 (SEQ ID NO:536), Figure 538 (SEQ ID NO:538), Figure 540 (SEQ ID NO:540), Figure 542 (SEQ ID NO:542), Figure 544 (SEQ ID NO:544), Figure 546 (SEQ ID NO:546), Figure 548 (SEQ ID NO:548) or Figure 550 (SEQ ID NO:550), lacking its associated signal peptide.

20. An isolated polypeptide having at least 80% amino acid sequence identity to:

(a) an amino acid sequence of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42),

NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312),

Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372), Figure 374 (SEQ ID NO:374), Figure 376 (SEQ ID NO:376), Figure 378 (SEQ ID NO:378), Figure 380 (SEQ ID NO:380), Figure 382 (SEQ ID NO:382), Figure 384 (SEQ ID NO:384), Figure 386 (SEQ ID NO:386), Figure 388 (SEQ ID NO:388), Figure 390 (SEQ ID NO:390), Figure 392 (SEQ ID NO:392), Figure 394 (SEQ ID NO:394), Figure 396 (SEQ ID NO:396), Figure 398 (SEQ ID NO:398), Figure 400 (SEQ ID NO:400), Figure 402 (SEQ ID NO:402), Figure 404 (SEQ ID NO:404), Figure 406 (SEQ ID NO:406), Figure 408 (SEQ ID NO:408), Figure 410 (SEQ ID NO:410), Figure 412 (SEQ ID NO:412), Figure 414 (SEQ ID NO:414), Figure 416 (SEQ ID NO:416), Figure 418 (SEQ ID NO:418), Figure 420 (SEQ ID NO:420), Figure 422 (SEQ ID NO:422), Figure 424 (SEQ ID NO:424), Figure 426 (SEQ ID NO:426), Figure 428 (SEQ ID NO:428), Figure 430 (SEQ ID NO:430), Figure 432 (SEQ ID NO:432), Figure 434 (SEQ ID NO:434), Figure 436 (SEQ ID NO:436), Figure 438 (SEQ ID NO:438), Figure 440 (SEQ ID NO:440), Figure 442 (SEQ ID NO:442), Figure 444 (SEQ ID NO:444), Figure 446 (SEQ ID NO:446), Figure 448 (SEQ ID NO:448), Figure 450 (SEQ ID NO:450), Figure 452 (SEQ ID NO:452), Figure 454 (SEQ ID NO:454), Figure 456 (SEQ ID NO:456), Figure 458 (SEQ ID NO:458), Figure 460 (SEQ ID NO:460), Figure 462 (SEQ ID NO:462), Figure 464 (SEQ ID NO:464), Figure 466 (SEQ ID NO:466), Figure 468 (SEQ ID NO:468), Figure 470 (SEQ ID NO:470), Figure 472 (SEQ ID NO:472), Figure 474 (SEQ ID NO:474), Figure 476 (SEQ ID NO:476), Figure 478 (SEQ ID NO:478), Figure 480 (SEQ ID NO:480), Figure 482 (SEQ ID NO:482), Figure 484 (SEQ ID NO:484), Figure 486 (SEQ ID NO:486), Figure 488 (SEQ ID NO:488), Figure 490 (SEQ ID NO:490), Figure 492 (SEQ ID NO:492), Figure 494 (SEQ ID NO:494), Figure 496 (SEQ ID NO:496), Figure 498 (SEQ ID NO:498), Figure 500 (SEQ ID NO:500), Figure 502 (SEQ ID NO:502), Figure 504 (SEQ ID NO:504), Figure 506 (SEQ ID NO:506), Figure 508 (SEQ ID NO:508), Figure 510 (SEQ ID NO:510), Figure 512 (SEQ ID NO:512), Figure 514 (SEQ ID NO:514), Figure 516 (SEQ ID NO:516), Figure 518 (SEQ ID NO:518), Figure 520 (SEQ ID NO:520), Figure 522 (SEQ ID NO:522), Figure 524 (SEQ ID NO:524), Figure 526 (SEQ ID NO:526), Figure 528 (SEQ ID NO:528), Figure 530 (SEQ ID NO:530), Figure 532 (SEQ ID NO:532), Figure 534 (SEQ ID NO:534), Figure 536 (SEQ ID NO:536), Figure 538 (SEQ ID NO:538), Figure 540 (SEQ ID NO:540), Figure 542 (SEQ ID NO:542), Figure 544 (SEQ ID NO:544), Figure 546 (SEQ ID NO:546), Figure 548 (SEQ ID NO:548) or Figure 550 (SEQ ID NO:550), lacking its associated signal peptide;

(b) an amino acid sequence of an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ

ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID

NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372), Figure 374 (SEQ ID NO:374), Figure 376 (SEQ ID NO:376), Figure 378 (SEQ ID NO:378), Figure 380 (SEQ ID NO:380), Figure 382 (SEQ ID NO:382), Figure 384 (SEQ ID NO:384), Figure 386 (SEQ ID NO:386), Figure 388 (SEQ ID NO:388), Figure 390 (SEQ ID NO:390), Figure 392 (SEQ ID NO:392), Figure 394 (SEQ ID NO:394), Figure 396 (SEQ ID NO:396), Figure 398 (SEQ ID NO:398), Figure 400 (SEQ ID NO:400), Figure 402 (SEQ ID NO:402), Figure 404 (SEQ ID NO:404), Figure 406 (SEQ ID NO:406), Figure 408 (SEQ ID NO:408), Figure 410 (SEQ ID NO:410), Figure 412 (SEQ ID NO:412), Figure 414 (SEQ ID NO:414), Figure 416 (SEQ ID NO:416), Figure 418 (SEQ ID NO:418), Figure 420 (SEQ ID NO:420), Figure 422 (SEQ ID NO:422), Figure 424 (SEQ ID NO:424), Figure 426 (SEQ ID NO:426), Figure 428 (SEQ ID NO:428), Figure 430 (SEQ ID NO:430), Figure 432 (SEQ ID NO:432), Figure 434 (SEQ ID NO:434), Figure 436 (SEQ ID NO:436), Figure 438 (SEQ ID NO:438), Figure 440 (SEQ ID NO:440), Figure 442 (SEQ ID NO:442), Figure 444 (SEQ ID NO:444), Figure 446 (SEQ ID NO:446), Figure 448 (SEQ ID NO:448), Figure 450 (SEQ ID NO:450), Figure 452 (SEQ ID NO:452), Figure 454 (SEQ ID NO:454), Figure 456 (SEQ ID NO:456), Figure 458 (SEQ ID NO:458), Figure 460 (SEQ ID NO:460), Figure 462 (SEQ ID NO:462), Figure 464 (SEQ ID NO:464), Figure 466 (SEQ ID NO:466), Figure 468 (SEQ ID NO:468), Figure 470 (SEQ ID NO:470), Figure 472 (SEQ ID NO:472), Figure 474 (SEQ ID NO:474), Figure 476 (SEQ ID NO:476), Figure 478 (SEQ ID NO:478), Figure 480 (SEQ ID NO:480), Figure 482 (SEQ ID NO:482), Figure 484 (SEQ ID NO:484), Figure 486 (SEQ ID NO:486), Figure 488 (SEQ ID NO:488), Figure 490 (SEQ ID NO:490), Figure 492 (SEQ ID NO:492), Figure 494 (SEQ ID NO:494), Figure 496 (SEQ ID NO:496), Figure 498 (SEQ ID NO:498), Figure 500 (SEQ ID NO:500), Figure 502 (SEQ ID NO:502), Figure 504 (SEQ ID NO:504), Figure 506 (SEQ ID NO:506), Figure 508 (SEQ ID NO:508), Figure 510 (SEQ ID NO:510), Figure 512 (SEQ ID NO:512), Figure 514 (SEQ ID NO:514), Figure 516 (SEQ ID NO:516), Figure 518 (SEQ ID NO:518), Figure 520 (SEQ ID NO:520), Figure 522 (SEQ ID NO:522), Figure 524 (SEQ ID NO:524), Figure 526 (SEQ ID NO:526), Figure 528 (SEQ ID NO:528), Figure 530 (SEQ ID NO:530), Figure 532 (SEQ ID NO:532), Figure 534 (SEQ ID NO:534), Figure 536 (SEQ ID NO:536), Figure 538 (SEQ ID NO:538), Figure 540 (SEQ ID NO:540), Figure 542 (SEQ ID NO:542), Figure 544 (SEQ ID NO:544), Figure 546 (SEQ ID NO:546), Figure 548 (SEQ ID NO:548) or Figure 550 (SEQ ID NO:550), with its associated signal peptide; or

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure

270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372), Figure 374 (SEQ ID NO:374), Figure 376 (SEQ ID NO:376), Figure 378 (SEQ ID NO:378), Figure 380 (SEQ ID NO:380), Figure 382 (SEQ ID NO:382), Figure 384 (SEQ ID NO:384), Figure 386 (SEQ ID NO:386), Figure 388 (SEQ ID NO:388), Figure 390 (SEQ ID NO:390), Figure 392 (SEQ ID NO:392), Figure 394 (SEQ ID NO:394), Figure 396 (SEQ ID NO:396), Figure 398 (SEQ ID NO:398), Figure 400 (SEQ ID NO:400), Figure 402 (SEQ ID NO:402), Figure 404 (SEQ ID NO:404), Figure 406 (SEQ ID NO:406), Figure 408 (SEQ ID NO:408), Figure 410 (SEQ ID NO:410), Figure 412 (SEQ ID NO:412), Figure 414 (SEQ ID NO:414), Figure 416 (SEQ ID NO:416), Figure 418 (SEQ ID NO:418), Figure 420 (SEQ ID NO:420), Figure 422 (SEQ ID NO:422), Figure 424 (SEQ ID NO:424), Figure 426 (SEQ ID NO:426), Figure 428 (SEQ ID NO:428), Figure 430 (SEQ ID NO:430), Figure 432 (SEQ ID NO:432), Figure 434 (SEQ ID NO:434), Figure 436 (SEQ ID NO:436), Figure 438 (SEQ ID NO:438), Figure 440 (SEQ ID NO:440), Figure 442 (SEQ ID NO:442), Figure 444 (SEQ ID NO:444), Figure 446 (SEQ ID NO:446), Figure 448 (SEQ ID NO:448), Figure 450 (SEQ ID NO:450), Figure 452 (SEQ ID NO:452), Figure 454 (SEQ ID NO:454), Figure 456 (SEQ ID NO:456), Figure 458 (SEQ ID NO:458), Figure 460 (SEQ ID NO:460), Figure 462 (SEQ ID NO:462), Figure 464 (SEQ ID NO:464), Figure 466 (SEQ ID NO:466), Figure 468 (SEQ ID NO:468), Figure 470 (SEQ ID NO:470), Figure 472 (SEQ ID NO:472), Figure 474 (SEQ ID NO:474), Figure 476 (SEQ ID NO:476), Figure 478 (SEQ ID NO:478), Figure 480 (SEQ ID NO:480), Figure 482 (SEQ ID NO:482), Figure 484 (SEQ ID NO:484), Figure 486 (SEQ ID NO:486), Figure 488 (SEQ ID NO:488), Figure 490 (SEQ ID NO:490), Figure 492 (SEQ ID NO:492), Figure 494 (SEQ ID NO:494), Figure 496 (SEQ ID NO:496), Figure 498 (SEQ ID NO:498), Figure 500 (SEQ ID NO:500), Figure 502 (SEQ ID NO:502), Figure 504 (SEQ ID NO:504), Figure 506 (SEQ ID NO:506), Figure 508 (SEQ ID NO:508), Figure 510 (SEQ ID NO:510), Figure 512 (SEQ ID NO:512), Figure 514 (SEQ ID NO:514), Figure 516 (SEQ ID NO:516), Figure 518 (SEQ ID NO:518), Figure 520 (SEQ ID NO:520), Figure 522 (SEQ ID NO:522), Figure 524 (SEQ ID NO:524), Figure 526 (SEQ ID NO:526), Figure 528 (SEQ ID NO:528), Figure 530 (SEQ ID NO:530), Figure 532 (SEQ ID NO:532), Figure 534 (SEQ ID NO:534), Figure

536 (SEQ ID NO:536), Figure 538 (SEQ ID NO:538), Figure 540 (SEQ ID NO:540), Figure 542 (SEQ ID NO:542), Figure 544 (SEQ ID NO:544), Figure 546 (SEQ ID NO:546), Figure 548 (SEQ ID NO:548) or Figure 550 (SEQ ID NO:550), lacking its associated signal peptide.

21. A method of detecting a PRO1801 polypeptide in a sample suspected of containing a PRO1801 polypeptide, said method comprising contacting said sample with a PRO1114 or PRO4978 polypeptide and determining the formation of a PRO1801/PRO1114 or PRO1801/PRO4978 polypeptide conjugate in said sample, wherein the formation of said conjugate is indicative of the presence of a PRO1801 polypeptide in said sample.

22. The method according to Claim 21, wherein said sample comprises cells suspected of expressing said PRO1801 polypeptide.

23. The method according to Claim 21, wherein said PRO1114 or PRO4978 polypeptide is labeled with a detectable label.

24. The method according to Claim 21, wherein said PRO1114 or PRO4978 polypeptide is attached to a solid support.

25. A method of detecting a PRO1114 or PRO4978 polypeptide in a sample suspected of containing a PRO1114 or PRO4978 polypeptide, said method comprising contacting said sample with a PRO1801 polypeptide and determining the formation of a PRO1801/PRO1114 or PRO1801/PRO4978 polypeptide conjugate in said sample, wherein the formation of said conjugate is indicative of the presence of a PRO1114 or PRO4978 polypeptide in said sample.

26. The method according to Claim 25, wherein said sample comprises cells suspected of expressing said PRO1114 or PRO4978 polypeptide.

27. The method according to Claim 25, wherein said PRO1801 polypeptide is labeled with a detectable label.

28. The method according to Claim 25, wherein said PRO1801 polypeptide is attached to a solid support.

29. A method of linking a bioactive molecule to a cell expressing a PRO1801 polypeptide, said method comprising contacting said cell with a PRO1114 or PRO4978 polypeptide that is bound to said bioactive molecule and allowing said PRO1801 and said PRO1114 or PRO4978 polypeptides to bind to one another, thereby linking said bioactive molecules to said cell.

30. The method according to Claim 29, wherein said bioactive molecule is a toxin, a radiolabel or an antibody.

31. The method according to Claim 29, wherein said bioactive molecule causes the death of said cell.

32. A method of linking a bioactive molecule to a cell expressing a PRO1114 or PRO4978 polypeptide, said method comprising contacting said cell with a PRO1801 polypeptide that is bound to said bioactive molecule and allowing said PRO1801 and said PRO1114 or PRO4978 polypeptides to bind to one another, thereby linking said bioactive molecules to said cell.

33. The method according to Claim 32, wherein said bioactive molecule is a toxin, a radiolabel or an antibody.

34. The method according to Claim 32, wherein said bioactive molecule causes the death of said cell.

35. A method of modulating at least one biological activity of a cell expressing a PRO1801 polypeptide, said method comprising contacting said cell with a PRO1114 or PRO4978 polypeptide or an anti-PRO1801 polypeptide antibody, whereby said PRO1114 or PRO4978 polypeptide or anti-PRO1801 polypeptide antibody binds to said PRO1801 polypeptide, thereby modulating at least one biological activity of said cell.

36. The method according to Claim 35, wherein said cell is killed.

37. A method of modulating at least one biological activity of a cell expressing a PRO1114 or PRO4978 polypeptide, said method comprising contacting said cell with a PRO1801 polypeptide or an anti-PRO1114 or anti-PRO4978 polypeptide antibody, whereby said PRO1801 polypeptide or anti-PRO1114 or anti-PRO4978 polypeptide antibody binds to said PRO1114 or PRO4978 polypeptide, thereby modulating at least one biological activity of said cell.

38. The method according to Claim 37, wherein said cell is killed.

39. A method of detecting a PRO1114 polypeptide in a sample suspected of containing a PRO1114 polypeptide, said method comprising contacting said sample with a PRO100 polypeptide and determining the formation of a PRO100/PRO1114 polypeptide conjugate in said sample, wherein the formation of said conjugate is indicative of the presence of a PRO1114 polypeptide in said sample.

40. The method according to Claim 39, wherein said sample comprises cells suspected of expressing said PRO1114 polypeptide.

41. The method according to Claim 39, wherein said PRO100 polypeptide is labeled with a detectable label.

5

42. The method according to Claim 39, wherein said PRO100 polypeptide is attached to a solid support.

10

43. A method of detecting a PRO100 polypeptide in a sample suspected of containing a PRO100 polypeptide, said method comprising contacting said sample with a PRO1114 polypeptide and determining the formation of a PRO100/PRO1114 polypeptide conjugate in said sample, wherein the formation of said conjugate is indicative of the presence of a PRO100 polypeptide in said sample.

15

44. The method according to Claim 43, wherein said sample comprises cells suspected of expressing said PRO100 polypeptide.

45. The method according to Claim 43, wherein said PRO1114 polypeptide is labeled with a detectable label.

20

46. The method according to Claim 43, wherein said PRO1114 polypeptide is attached to a solid support.

25

47. A method of linking a bioactive molecule to a cell expressing a PRO100 polypeptide, said method comprising contacting said cell with a PRO1114 polypeptide that is bound to said bioactive molecule and allowing said PRO100 and said PRO1114 polypeptides to bind to one another, thereby linking said bioactive molecules to said cell.

30

48. The method according to Claim 47, wherein said bioactive molecule is a toxin, a radiolabel or an antibody.

49. The method according to Claim 47, wherein said bioactive molecule causes the death of said cell.

35

50. A method of linking a bioactive molecule to a cell expressing a PRO1114 polypeptide, said method comprising contacting said cell with a PRO100 polypeptide that is bound to said bioactive molecule and allowing said PRO100 and said PRO1114 polypeptides to bind to one another, thereby linking said bioactive molecules to said cell.

51. The method according to Claim 50, wherein said bioactive molecule is a toxin, a radiolabel or an antibody.

52. The method according to Claim 50, wherein said bioactive molecule causes the death of said cell.

53. A method of modulating at least one biological activity of a cell expressing a PRO100 polypeptide, said method comprising contacting said cell with a PRO1114 polypeptide or an anti-PRO100 polypeptide antibody, whereby said PRO1114 polypeptide or anti-PRO100 polypeptide antibody binds to said PRO100 polypeptide, thereby modulating at least one biological activity of said cell.

54. The method according to Claim 53, wherein said cell is killed.

55. A method of modulating at least one biological activity of a cell expressing a PRO1114 polypeptide, said method comprising contacting said cell with a PRO100 polypeptide or an anti-PRO1114 polypeptide antibody, whereby said PRO100 polypeptide or anti-PRO1114 polypeptide antibody binds to said PRO1114 polypeptide, thereby modulating at least one biological activity of said cell.

56. The method according to Claim 55, wherein said cell is killed.

57. A method for stimulating the release of TNF- α from human blood, said method comprising contacting said blood with a PRO195, PRO202, PRO215, PRO221, PRO217, PRO222, PRO198, PRO245, PRO172, PRO265, PRO266, PRO344, PRO337, PRO322, PRO1286, PRO1279, PRO1338 or PRO1343 polypeptide, wherein the release of TNF- α from said blood is stimulated.

58. A method for modulating the uptake of glucose or FFA by skeletal muscle cells, said method comprising contacting said cells with a PRO182, PRO366, PRO198, PRO172 or PRO719 polypeptide, wherein the uptake of glucose or FFA by said cells is modulated.

59. A method for stimulating the proliferation or differentiation of chondrocyte cells, said method comprising contacting said cells with a PRO182, PRO366, PRO198, PRO1868, PRO202, PRO224, PRO172, PRO301 or PRO1312 polypeptide, wherein the proliferation or differentiation of said cells is stimulated.

60. A method for modulating the uptake of glucose or FFA by adipocyte cells, said method comprising contacting said cells with a PRO202, PRO211, PRO344 or PRO1338 polypeptide, wherein the uptake of glucose or FFA by said cells is modulated.

61. A method for stimulating the proliferation of or gene expression in pericyte cells, said method comprising contacting said cells with a PRO366 polypeptide, wherein the proliferation of or gene expression in said cells is stimulated.

5 62. A method for stimulating the release of proteoglycans from cartilage, said method comprising contacting said cartilage with a PRO216 polypeptide, wherein the release of proteoglycans from said cartilage is stimulated.

63. A method for stimulating the proliferation of inner ear utricular supporting cells, said method comprising contacting said cells with a PRO172 polypeptide, wherein the proliferation of said cells is stimulated.

10 64. A method for stimulating the proliferation of T-lymphocyte cells, said method comprising contacting said cells with a PRO344 polypeptide, wherein the proliferation of said cells is stimulated.

15 65. A method for stimulating the release of a cytokine from PBMC cells, said method comprising contacting said cells with a PRO526 or PRO1343 polypeptide, wherein the release of a cytokine from said cells is stimulated.

20 66. A method for inhibiting the binding of A-peptide to factor VIIA, said method comprising contacting a composition comprising said A-peptide and said factor VIIA with a PRO182 polypeptide, wherein the binding of said A-peptide to said factor VIIA is inhibited.

67. A method for inhibiting the differentiation of adipocyte cells, said method comprising contacting said cells with a PRO185 or PRO198 polypeptide, wherein the differentiation of said cells is inhibited.

25 68. A method for stimulating the proliferation of endothelial cells, said method comprising contacting said cells with a PRO222 polypeptide, wherein the proliferation of said cells is inhibited.

30 69. A method for detecting the presence of tumor in an mammal, said method comprising comparing the level of expression of any PRO polypeptide shown in Table 8 in (a) a test sample of cells taken from said mammal and (b) a control sample of normal cells of the same cell type, wherein a higher level of expression of said PRO polypeptide in the test sample as compared to the control sample is indicative of the presence of tumor in said mammal.

35 70. The method of Claim 69, wherein said tumor is lung tumor, colon tumor, breast tumor, prostate tumor, rectal tumor, cervical tumor or liver tumor.

71. An oligonucleotide probe derived from any of the nucleotide sequences shown in the accompanying figures.

1/550

FIGURE 1

GTTACTCGGTGGTGGCGGAGTCTACGGAAGCCGTTTTTCGCTTCACTTTTCCTGGCTGTAGAGC
GCTTTCCCCCTGGCGGGTGAGAGTGCAGAGACGAAGGTGCGAGATGAGCACTATGTTGCGCGGA
CACTCTCCTCATCGTTTTTATCTCTGTGTGCACGGCTCTGCTCGCAGAGGGCATAACCTGGGT
CCTGGTTTACAGGACAGACAAGTACAAGAGACTGAAGGCAGAAGTGGAACAGAGTAAAAA
ATTGGAAGAAGAAGGAAACAATAACAGAGTCAGCTGGTCGACAACAGAAAAAGAAAATAGA
GAGACAAGAAGAGAACTGAAGAATAACAACAGAGATCTATCAATGGTTTGAATGAAATCCAT
GTTTGCTATTGGCTTTTGTCTTACTGCCCTAATGGGAATGTTCAATTCCATATTTGATGGTAG
AGTGGTGGCAAAGCTTCCTTTTACCCCTCTTTCTTACATCCAAGGACTGTCTCATCGAAATCT
GCTGGGAGATGACACCACAGACTGTTCCCTTCATTTTCCTGTATATTCTCTGTACTATGTGAT
TCGACAGAACATTGAGAAGATTCTCGGCCTTGCCCTTCACGAGCCGCCACCAAGCAGGCAGG
TGGATTTCTTGCCCAACACCTCCTTCTGGGAAGTTCTCTTGAACTCAAGAACTCTTTATTTT
CTATCATTCTTTCTAGACACACACATCAGACTGGCAACTGTTTTGTAGCAAGAGCCATAGG
TAGCCTTACTACTTGGGCCTCTTTCTAGTTTTGAATTATTTCTAAGCCTTTTGGGTATGATTA
GAGTGAAAATGGCAGCCAGCAAACCTTGATAGTGCTTTTGGTCCTAGATGATTTTTATCAAATA
AGTGGATTGATTAGTTAAGTTCAGGTAATGTTTATGTAATGAAAAACAAATAGCATCCTTCTT
GTTTCATTTACATAAGTATTTTCTGTGGGACCGACTCTCAAGGCACTGTGTATGCCCTGCAAG
TTGGCTGTCTATGAGCATTTAGAGATTTAGAAGAAAAATTTAGTTTGTTTAACCCTTGTAAC
GTTTGTTTTGTTGTTGTTTTTTTTTCAAGCCAAATACATGACATAAGATCAATAAAGAGGCCA
AATTTTTAGCTGTTTTATGTACAAGGAGAGATCTGTTTCATTTTGTTTTGCCGTATTTCTAGA
TATAAGTTTTAGCATGGGCCAGGAAGGACTAAAATAAAAGTTTTTAAGGTACAAAAAAAAAAAA
AAAA

2/550

FIGURE 2

MSTMFADTLLIVFISVCTALLAEGITWVLVYRTDKYKRLKAEVEKQSKKLEKKKETITESAGR
QQKKKIERQEEKLKNNNRDLSMVRMKSMFAIGFCFTALMGMFNSIFDGRVVAKLPFTPLSYIQ
GLSHRNLLGDDTTDCSFIFLYIILCTMSIRQNIQKILGLAPSRAATKQAGGFLGPPPPSGKFS

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites..

amino acids 103-109, 163-169

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 53-57

3/550

FIGURE 3

AGCCGGGGGCGGGTTTGAAGACGCGTCGTTGGGTTTGGAGGCCGTGAAACAGCCGTTTGAGT
TTGGCTGCGGGTGGAGAACGTTTGTGAGGGGCCCGGCCAAGAAGGAGGCCCGCCTGTTACGAT
GGTGTCCATGAGTTTCAAGCGGAACCGCAGTGACCGGTTCTACAGCACCCGGTGCTGCGGCTG
TTGCCATGTCCGCACCGGGACGATCATCCTGGGGACCTGGTACATGGTAGTAAACCTATTGAT
GGCAATTTTGCTGACTGTGGAAGTGACTCATCCAACTCCATGCCAGCTGTCAACATTCAGTA
TGAAGTCATCGGTAATTACTATTGCTCTGAGAGAATGGCTGATAATGCCTGTGTTCTTTTGC
CGTCTCTGTTCTTATGTTTATAATCAGTTCAATGCTGGTTTATGGAGCAATTTCTTATCAAGT
GGGTTGGCTGATTCCATTCTTCTGTTACCGACTTTTTGACTTCGTCCTCAGTTGCCTGGTTGC
TATTAGTTCTCTCACCTATTTGCCAAGAATCAAAGAATATCTGGATCAACTACCTGATTTTCC
CTACAAAGATGACCTCCTGGCCTTGGACTCCAGCTGCCTCCTGTTTATTGTTCTTGTGTTCTT
TGCCTTATTCATCATTTTTTAAGGCTTATCTAATTAAGTGTGTTTGGAACTGCTATAAATACAT
CAACAACCGAAACGTGCCGGAGATTGCTGTGTACCCTGCCTTTGAAAGCACCTCCTCAGTACG
TTTTGCCAACCTATGAAATGGCCGTGAAAATGCCTGAAAAAGAACCACCACCTCCTTACTTAC
CTGCCTGAAGAAATTCTGCCTTTGACAATAAATCCTATACCAGCTTTTTGTTTGTTTATGTTA
CAGAATGCTGCAATTCAGGGCTCTTCAAACCTGTTTGATATAAAATATGTTGTCTTTTGTTA
AGCATTTATTTCAAACACTAAGGAGCTTTTTGACATCTGTTAAACGTCTTTTTGTTTTTTG
TTAAGTCTTTTACATTTTAATAGTTTTTGAAGACAATCTAGGTAAAGCAAGAGCAAAGTGCCA
TTGTTTGCCTTTAATTGGGGGGTGGGAAGGGAAGAGGGTACTTGCCACATAGTTTCCTTTTT
AACTGCACTTTCTTTATATAATCGTTTGCATTTTGTACTTGCTACCCTGAGTACTTTCAGGA
AGACTGACTTAAATATTCGGGGTGAGTAAGTAGTTGGGTATAAGATCTGAACTTTTCATCTGC
AGAGGCAAGAAAAATATTTGACATTGTGACTTGACTGTGGAAGATGATGGTTGCATGTTTCTA
GTTTGTATATGTTTCCATCTTTGTGATAAGATGATTTAATAAATCTCTTTAAATACTAAAAA
AAAAAAA

4/550

FIGURE 4

MVSMSFKRNRSDRFYSTRCCGCCHVRTGTIILGTWYMVVNLLMAILLTVEVTHPNSMPAVNIQ
YEVIGNYYSSERMADNACVLFAVSVLMEFISSMLVYGAIYQVGWLIPFFCYRLFDFVLSCLV
AISSLTYLPRIKEYLDQLPDFPYKDDLLALDSSCLLFIVLVFFALFIIFKAYLINCWVWNCYKY
INNRRNVPEIAVYPAFESTSSVRFANL

Important features of the protein:**Transmembrane domain (Possible type II transmembrane protein):**

amino acids 30-49, 81-100, 111-131, 158-175

N-glycosylation site.

amino acids 9-13

Tyrosine kinase phosphorylation sites.

amino acids 8-16, 193-202

N-myristoylation site.

amino acids 68-74

5/550

FIGURE 5

CCCGCTGGCCCGTCAGTGCTCTCCCCGTCGTTTGCCCTCTCCAGTTCCCCCAGTGCCTGCCCT
ACGCACCCCGATGGCGGAGCTGCGGCCTAGCGGCGCCCCGGCCCCACCGCGCCCCCGGCCCC
TGGCCCGACTGCCCCCGGCCCTTCGCTTCGCTCTTTCCCCCGGGACTGCACGCCATCTACGG
AGAGTGCCGCGCCTTTACCCTGACCAGCCGAACCCGCTCCAGGTTACCGCTATCGTCAAGTA
CTGGTTGGGTGGCCCAGACCCCTTGGACTATGTTAGCATGTACAGGAATGTGGGGAGCCCTTC
TGCTAACATCCCCGAGCACTGGCACTACATCAGCTTCGGCCTGAGTGATCTCTATGGTGACAA
CAGAGTCCATGAGTTTACAGGAACAGATGGACCTAGTGGTTTTGGCTTTGAGTTGACCTTTCG
TCTGAAGAGAGAACTGGGGAGTCTGCCCCACCAACATGGCCCGCAGAGTTAATGCAGGGCTT
GGCACGATACGTGTTCCAGTCAGAGAACACCTTCTGCAGTGGGGACCATGTGTCCTGGCACAG
CCCTTTGGATAACAGTGAGTCAAGAATTGACACATGCTGCTGACAGAGGACCCACAGATGCA
GCCCGTGCAGACACCCTTTGGGGTAGTTACCTTCCTCCAGATCGTTGGTGTCTGCACTGAAGA
GCTACACTCAGCCAGCAGTGGAACGGGCAGGGCATCCTGGAGCTGCTGCGGACAGTGCCTAT
TGCTGGCGGGCCCTGGCTGATAACTGACATGCGGAGGGGAGAGACCATATTTGAGATCGATCC
ACACCTGCAAGAGAGAGTTGACAAAGGCATCGAGACAGATGGCTCCAACCTGAGTGGTGTGAG
TGCCAAGTGTGCCTGGGATGACCTGAGCCGGCCCCCGAGGATGACGAGGACAGCCGGAGCAT
CTGCATCGGCACACAGCCCCGGCGACTCTCTGGCAAAGACACAGAGCAGATCCGGGAGACCCT
GAGGAGAGGACTCGAGATCAACAGCAAACCTGTCTTCCACCAATCAACCCTCAGCGGCAGAA
TGGCCTCGCCCACGACCGGGCCCCGAGCCGCAAAGACAGCCTGGAAAGTGACAGCTCCACGGC
CATCATTTCCCATGAGCTGATTTCGCACGCGGCAGCTTGAGAGCGTACATCTGAAATTC AACCA
GGAGTCCGGAGCCCTCATTCCTCTCTGCCTAAGGGGCAGGCTCCTGCATGGACGGCACTTTAC
ATATAAAAGTATCACAGGTGACATGGCCATCACGTTTGTCTCCACGGGAGTGGAAGGCGCCTT
TGCCACTGAGGAGCATCCTTACGCGGCTCATGGACCCTGGTTACA ACTCTGAACCTATCCTCG
GAGCTCTGCCCTCCCGTCTTGAACGTCTTTCTGCCCTGAGGAGAGGGTAGTCAGCATCTCCA
ATTTTCAGCAGCTCAAGAACCTTGGCCCCCACAGGACTTCGCAGATGTCACATTGCCCTCAG
TCCCCTGAATGCCCTTCGGACCCAACCCCAATTCCCCAAGCCCCTGACCCCTAGCTGCCGGG
GTTCCCCTCCCAGTGCCACAACCCCTCACCTCCCCTGGCAGCCCCTCAGCGAGCCTGAGGC
CCAGCACCCGCTGGCTCCCCAGCACATGGTCCCCTCCCATGGGCTGTTGCCAGGGAACCGGG
GCGCGGTGGGAACGAGCTGCTGGCCTCGGCATGTTTCAATAAAGTTGCTGTGCTGGGAG

6/550

FIGURE 6

MAELRPSGAPGPTAPPAPGPTAPPAFASLFPPGLHAIYGECRRLYPDQPNPLQVTAIVKYWLG
GPDPLDYVSMYRNVGSPSANIPEHWHYISFGLSDLYGDNRVHEFTGTDGPGSGFGFELTFRLKR
ETGESAPPTWPAELMQGLARYVFQSENTFCSGDHVSWHSPLDNSESRIQHMLLTEDPQMOPVQ
TPFGVVTFLQIVGVCTEELHSAQQWNGQGILELLRTVPIAGGPWLITDMRRGETIFEIDPHLQ
ERVDKGIETDGSNLSGVSAKCAWDDLSRPPEDDEDSRSICIGTQPRRLSGKDTEQIRETLRRG
LEINSKPVLPPIPNQRQNGLAHDRAPSRKDSLESSTAIIPHELIRTRQLESVHLKFNQESG
ALIPLCLRGRLLHGRHFTYKSITGDMAITFVSTGVEGAFATEEHPYAAHGFWLQL

Important features:**N-glycosylation site.**

amino acids 265-268

7/550

FIGURE 7

CGCGAATGAAGTTTGCATTTTCCTCTGTTCTTGAGCCCAGCTTCTTCTCGTCTCCACCCCCAG
CTTCCCGGCATTGGAAGAAGGGACCGTCCTCTTCCTTGTCTTGGCCACCCAAATCCTGGTATC
GAAAGGGTTGAACGGACCGGAAGTGTGCAGCAGCGACGGGTCCCCAGCTAATCGACGCCGGAA
GTAGCAATTACTAGACAAGCATTCCGCCGCCGGCTTCGCTATGGCGGCAATTCCCCCAGATTCT
CTGGCAGCCACCCAACGTTTACTTGGAGACCAGCATGGGAATCATTGTGCTGGAGCTGTACTG
GAAGCATGCTCCAAAGACCTGTAAGAACTTTGCTGAGTTGGCTCGTCGAGGTTACTACAATGG
CACAAAATTCCACAGAATTATCAAAGACTTCATGATCCAAGGAGGTGACCCAACAGGGACAGG
TCGAGGTGGTGCATCTATCTATGGCAAACAATTTGAAGATGAACTTCATCCAGACTTGAAATT
CACGGGGGCTGGAATTCTCGCAATGGCCAATGCGGGGCCAGATACCAATGGCAGCCAGTTCTT
TGTGACCCCTCGCCCCACCCAGTGGCTTGACGGCAAACACACCATTTTTGGCCGAGTGTGTCA
GGGCATAGGAATGGTGAATCGCGTGGGAATGGTAGAAACAACTCCCAGGACCGCCCTGTGGA
CGACGTGAAGATCATTAAGGCATACCCTTCTGGGTAGACTTGCTACCCTCTTGAGCAGCTCTT
CTGAGATGGCCCCAGTGAACCAGCTTCTAGATGACATAGAATGACATGTAATGCTAAATTTCA
TTTTGGCTTTGCAAGTCATGAAGCTTAGGAGGCCTGGCATCTTGGGTGAGTTAGAGATGGAAG
TACATTTTAATAGGATGCTTCTTTTCTCTTCCCCAGTGCCTAGGTTGCCAGAGCATTTCAC
AATGCCCTGTTTATCAATAGGTGACTACTTACTACACATGAACCATAATGCTGCTTCTTGT
GCATGTCTGCTCTGATATACGTCGAACAATGTAGCAGCCACTGTCATTTCTCAGTGGTTTTGC
CTAACCAAATCTTCTCTAAGGAGATTTATATTCTGGCCTACACAGCAGTCCTTGATGGCTGA
CAGCCACAGAATTCCAAACCAAGTAGTGTCTGTCAGCCCTCTTAACTCTGTGCACGCCCTATT
TCAGTCTTTTACATTTGTTCTTCTAGGGAATGTATGCATCTCTATATATATTTTCCCTCTCAA
AACCAGAACATCAACAGTGCTGTTTCTGACACTTCAGACATCCCACGCAAAGCCACATTGAAT
TTTTGCCAAATGAAAAACACATCCAACAATCAAGTTTCTAAGAAGGTGTCAAGTGGGGAATAA
TAATAATGTATAATAATCAAGAAATTAGTTTATTAAAGGAAGCAGAAGCATTGACCATTTTT
TCCCAGAGAAGAGGAGAAATCTGTAGTGAGCAAAGGACAGACCATGAATCCTCCTTGAGAAGT
AGTACTCTCAGAAAGGAGAAGCGCCACTCAAGTTCTTTTAACCAAGACTTTAGAGAAATTAG
GTCCAAGATTTTTATATGTTTCTGTTTATGTATAAAAATAACTTCTGGATTTTGTGGGGA
GGAGCAGGAGAGGAAGGAAGTTAATACCTATGTAATACATAGAACTTCCACAATAAAATGCC
ATTGATGGTTAAAAAAAAAAAAAAAAAAAA

8/550

FIGURE 8

MAAIPDSWQPPNVYLETSMGIIIVLELYWKHAPKTCKNFAELARRGYNGTKFHRIKDFMIQ
GGDPTGTGRGGASIYGKQFEDELHPDLKFTGAGILAMANAGPDTNGSQFFVTLAPTQWLDGKH
TIFGRVCQGIGMVNRVGMVETNSQDRPVDDVKI IKAYPSG

Important features:

N-glycosylation sites:

amino acids 49-52, 108-111

N-myristoylation sites:

amino acids 64-69, 69-74, 143-148

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature:

amino acids 48-65

9/550

FIGURE 9

[illegible]

10/550

FIGURE 10

MWHEARKHERKLRGMMVDYKKRAERRREYYEYKIKKDPAQFLQVHGRACKVHLDSAVALAAESP
VNMPWQGDNTNNMIDRFDVRAHLDHIPDYTPPLLTTISPEQESDERKCNYERYRGLVQNDFAFAG
ISEEQCLYQIYIDELYGGLQRPSEDEKKKLAEEKKASIGYTYEDSTVAEVEKAAEKPEEEESAA
EEESNSDEDEVIPDIDVEVDVDELNQEQQVADLNKQATTYGMADGDFVRLRDKKEEAEAIKHA
KALEEEKAMYSGRRSRRQRREFREKRLRGRKISPPSYARRDSPTYDPYKRSPSESSSESRSRS
RSPTPGREEKITFITSFGGSDEEAAAAAAAAAASGVTTGKPPAPPQPGGPAPGRNASARRRSS
SSSSSSSASRTSSSRSSSRSSSRSGGGYYRSGRHARSRSRSWSRSRSRSRRYSRSRSRGRR
HSGGGSRDGHYRSRSPARRGGYGPRRRSRRSRSHSGDRYRRGGGRGLRHHSSSRSSWSLSPSR
SRSLTRSRSHSPSPSQSRSRSRSPSPSPAREKLTRPAASPAVGEKLLKTEPAAGKETGA
AKVTQADASGEAETEDAEGAEQAVQGG

Important features:**N-glycosylation site:**

amino acids 370-373

Glycosaminoglycan attachment site:

amino acids 443-446

cAMP- and cGMP-dependent protein kinase phosphorylation site:amino acids 159-162, 282-285, 291-294, 374-377, 375-378, 430-433,
440-443, 466-469**Casein kinase II phosphorylation site:**amino acids 149-152, 166-169, 171-174, 187-190, 193-196, 195-198,
303-306, 307-310, 335-338, 571-574**N-myristoylation sites:**

amino acids 118-123, 229-234, 350-355, 446-451, 586-591

Amidation sites:

amino acids 263-266, 280-283, 438-441

11/550

FIGURE 11

GGTAGGCGCGCCAGACCTGAGACGGGTTGGGACTGGGCTGCGTCACGCGCGGGCTCTAAGCG
CCCGGGGCCCCGCCAGTGCCGGGCACAGCCAATCGCAGCGCGGAAGGCGGTGGGGGCGGGG
AAGGCCGCTTGAAACTTAAATCCCGAGGCGGGCGAACCTGCACCAGACCGCGGACGTCTGTA
ATCTCAGAGGCTTGTTTGCTGAGGGTGCCTGCGCAGCTGCGACGGCTGCTGGTTTTGAAACAT
GAATCTTTCGCTCGTCCTGGCTGCCTTTTGCTTGGGAATAGCCTCCGCTGTTCCAAAATTTGA
CCAAAATTTGGATACAAAGTGGTACCAGTGGAAGGCAACACACAGAAGATTATATGGCGCGAA
TGAAGAAGGATGGAGGAGAGCAGTGTGGGAAAAGAATATGAAAATGATTGAACTGCACAATGG
GGAATACAGCCAAGGGAAACATGGCTTCACAAATGGCCATGAATGCTTTTGGTGACATGACCAA
TGAAGAATTCAGGCAGATGATGGGTTGCTTTCGAAACCAGAAATTCAGGAAGGGGAAAGTGTT
CCGTGAGCCTCTGTTTCTTGATCTTCCCAAATCTGTGGATTGGAGAAAGAAAGGCTACGTGAC
GCCAGTGAAGAATCAGAAACAGTGTGGTTCTTGTTGGGCTTTTAGTGCGACTGGTGCTCTTGA
AGGACAGATGTTCCGGAAAACCTGGGAAACTTGTCTCACTGAGCGAGCAGAATCTGGTGGACTG
TTCGCGTCCTCAAGGCAATCAGGGCTGCAATGGTGGCTTCATGGCTAGGGCCTTCCAGTATGT
CAAGGAGAACGGAGGCCTGGACTCTGAGGAATCCTATCCATATGTAGCAGTGGATGAAATCTG
TAAGTACAGACCTGAGAATTCTGTTGCTAATGACACTGGCTTCACAGTGGTCGCACCTGGAAA
GGAGAAGGCCCTGATGAAAGCAGTCGCAACTGTGGGGCCCATCTCCGTTGCTATGGATGCAGG
CCATTCGTCTTCCAGTTCTACAAATCAGGCATTTATTTTGAACCAGACTGCAGCAGCAAAAA
CCTGGATCATGGTGTTCTGGTGGTTGGCTACGGCTTTGAAGGAGCAAATTCGAATAACAGCAA
GTATTGGCTCGTCAAAAACAGCTGGGGTCCAGAATGGGGCTCGAATGGCTATGTAAAAATAGC
CAAAGACAAGAACAACCACTGTGGAATCGCCACAGCAGCCAGCTACCCCAATGTGTCAGCTGA
TGGATGGTGAGGAGGAAGGACTTAAGGACAGCATGTCTGGGGAAATTTTATCTTGAACTGAC
CAAACGCTTATTGTGTAAGATAAACCACTTGAATCATGGAGGATCCAAGTTGAGATTTTAATT
CTGTGACATTTTTACAAGGGTAAAAATGTTACCACTACTTTAATTATTGTTATACACAGCTTTA
TGATATCAAAGACTCATTGCTTAATTCTAAGACTTTTGAATTTTCATTTTTTAAAAAGATGTA
CAAAACAGTTTGAAATAAATTTTAATTCGTATATA

12/550

FIGURE 12

MNLSLVLA AFCLGIASAVPKFDQNLDTKWYQWKATHRRLYGANEEGWRRRAVWEKNMKMIELHN
GEYSQKGHGFTMAMNAFGDMTNEEFRQMMGCFRNQKFRKGKVFREPLFLDLPKSV DWRKKGYV
TPVKNQKQCGSCWAFSATGALEGQMFRKTGKLVSLSEQNLVDCSRPQGNQGCNGGF MARAFQY
VKENGGLDSEESYPYVAVDEICKYRPENSVANDTGFTVVAPGKEKALMKAVATVGPISVAMDA
GHSSFQFYKSGIYFEPDCSSKNLDHGVLVVGYGFE GANSNNSKYWLVKNSWGPEWGSNGYVKI
AKDKNNHCGIATAASYPNV

Important features:**Signal sequence**

amino acids 1-17

N-glycosylation sites.

amino acids 2-6, 221-225, 292-296

N-myristoylation sites.amino acids 13-19, 93-99, 136-142, 145-151, 174-180, 177-183,
180-186, 194-200, 288-294, 324-330**Eukaryotic thiol (cysteine) proteases cysteine active site.**

amino acids 132-144

Eukaryotic thiol (cysteine) proteases histidine active site.

amino acids 275-286

13/550

FIGURE 13

GGCGGCGTCATGTGATCCGCTTCCCTGCTCCTTTAAGCGTCCACAGGCGGGGAGCGGCCACA
ATCACAGCTCCGGGCATTGGGGGAACCCGAGCCGGCTGCGCCGGGGGAATCCGTGCGGGCGCC
TTCCGTCCCGGTCCCATCCTCGCCGCGCTCCAGCACCTCTGAAGTTTTGCAGCGCCCAGAAAG
GAGGCGAGGAAGGAGGGAGTGTGTGAGAGGAGGGAGCAAAAAGCTCACCTAAAACATTTATT
TCAAGGAGAAAAGAAAAGGGGGGGCGCAAAAATGGCTGGGGCAATTATAGAAAACATGAGCA
CCAAGAAGCTGTGCATTGTTGGTGGGATTCTGCTCGTGTTCCAAATCATCGCCTTTCTGGTGG
GAGGCTTGATTGCTCCAGGGCCACAACGGCAGTGTCTACATGTGGTGAAATGTGTGGATG
CCCGTAAGAACCATCACAAAGCAAAATGGTTTCGTGCCTTGGGGACCCAATCATTGTGACAAGA
TCCGAGACATTGAAGAGGCAATTCCAAGGGAAATTGAAGCCAATGACATCGTGTTTTCTGTTC
ACATTCCCCCTCCCCACATGGAGATGAGTCCTTGGTTCCAATTCATGCTGTTTATCCTGCAGC
TGGACATTGCCTTCAAGCTAAACAACCAATCAGAGAAAATGCAGAAGTCTCCATGGACGTTT
CCCTGGCTTACCGTGATGACGCATTTGCTGAGTGGACTGAAATGGCCCATGAAAGAGTACCAC
GGAACTCAAATGCACCTTCACATCTCCCAAGACTCCAGAGCATGAGGGCCGTTACTATGAAT
GTGATGTCCTTCTTTTCATGGAAATTGGGTCTGTGGCCCATAGTTTTACCTTTTAAACATCC
GGCTGCCTGTGAATGAGAAGAAGAAAATCAATGTGGGAATTGGGGAGATAAAGGATATCCGGT
TGGTGGGGATCCACCAAAATGGAGGCTTCACCAAGGTGTGGTTGCCATGAAGACCTTCCTTA
CGCCCAGCATCTTCATCATTATGGTGTGGTATTGGAGGAGGATCACCATGATGTCCCGACCCC
CAGTGCTTCTGGAAAAAGTCATCTTTGCCCTTGGGATTTCCATGACCTTTATCAATATCCCAG
TGGAAATGGTTTTCCATCGGGTTTGACTGGACCTGGATGCTGCTGTTTGGTGACATCCGACAGG
GCATCTTCTATGCGATGCTTCTGTCTTCTGGATCATCTTCTGTGGCGAGCACATGATGGATC
AGCACGAGCGGAACACATCGCAGGGTATTGGAAGCAAGTCGGACCCATTGCCGTTGGCTCCT
TCTGCCTCTTCATATTTGACATGTGTGAGAGAGGGGTACAACCTCACGAATCCCTTCTACAGTA
TCTGGACTACAGACATTGGAACAGAGCTGGCCATGGCCTTCATCATCGTGGCTGGAATCTGCC
TCTGCCTCTACTTCTGTCTTCTATGCTTCATGGTATTTTCAGGTGTTTCGGAACATCAGTGGGA
AGCAGTCCAGCCTGCCAGCTATGAGCAAAGTCGGCGGCTACACTATGAGGGGCTAATTTTTTA
GGTTCAAGTTCCTCATGCTTATCACCTTGGCCTGCGCTGCCATGACTGTCATCTTCTTCATCG
TTAGTCAGGTAACGGAAGGCCATTGGAATGGGGCGGCGTCACAGTCCAAGTGAACAGTGCCT
TTTTACAGGCATCTATGGGATGTGGAATCTGTATGTCTTTGCTCTGATGTTCTTGTATGCAC
CATCCCATAAAACTATGGAGAAGACCAGTCCAATGGCGATCTGGGTGTCCATAGTGGGGGAAG
AACTCCAGCTCACCACCACTATCACCCATGTGGACGGACCCACTGAGATCTACAAGTTGACCC
GCAAGGAGGCCCAGGAGTAGGAGGCTGCAGCGCCCGGCTGGGACGGTCTCTCCATACCCAGC
CCCTCTAACTAGAGTGGGGAGCATGCCAGAGAGAGCTCAATGTACAAATGAATGCCTCATGGC
TCTTAGCTGTGGTTTTCTTGGACCAGCGGCATGGACATTTGTGAGTTTGCTTCTGACGGTAGC
TTTTGGAGGAAGATTCTTGACGCCACTAATGCATTGTGTATGATAACAAAACCTCTGGTATGA
CACATTTTCTGTGATCATTGTTAATTAGTGACATAGTAACATCTGTAGCAGCTGGTTAGTAAA
CCTCATGTGGGGGTGGGGTGGGGGTGTATTCTTGGGGGATGGTTTGGGCCGAATGGGGAGTG
GAATATTTGACATTTTCTGTGTTTTAAATCTAGGATAGATTTTAAACATCCTTTGCGGTCCCA
GTCCAAGGTAGGCTGGTGTATAGTCTTCTCACTCCTAATCCATGACCACTGTTTTTTTCTTA
TTTATATCACCAGGTAGCCTACTGAGTTAATATTTAAGTTGTCAATAGATAAGTGTCCCTGTT
TTGTGGCATAATAAAGTGAATTTTCATGAGAAGATTTTATCCACCAGGGGTATTTTCAAGTTG
AAACCAATCTGTGTATCTAATACTAACCAATCTGTTGGATGTGGATTTTAAAAAAGTTTTC
TAACTACCCAAGTAAGATTTACTGTATTAATGGCCTTCGGGTCTGAAAAGCTTTTTTAAACC
TCTTGCTTAAATGCGTTTTTATTTTGATAAGATACTTCAAATAGCCTCCAAAAGTGTAGATCC
AATCACTTAAATAAACCTGTATGTATATGCAAAAAAAAAAAAAAAAAAAAAA

14/550

FIGURE 14

MAGAI IENMSTKKLCIVGGILLVFQIIAFLVGGLIAPGPTTAVSYMSVKCVDARKNHHKTKWF
VPWGNHCDKIRDIEEAI PREIEANDIVFSVHIPLPHMEMSPWFQFMLFILQLDIAFKLNNQI
RENAEVSMDSLAYRDDAFAEWTEMAHERVPRKLKCTFTSPKTPEHEGRYYECDVLPFMEIGS
VAHKFYLLNIRLPVNEKKKINVGIGEIKDIRLVGIHQNGGFTKVWFAMKTFLTSPSIFIIMVWY
WRRITMSRPPVLLEKVI FALGISMTFINIPVEWFSIGFDWTWMLLFGDIRQGIFYAMLLSEW
IIFCGEHMMDQHERNHIAGYWKQVGPIAVGSFCLFIFDMCERGVQLTNPFYSIWTTDIGTELA
MAFIIVAGICLCLYFLFLCFMVFQVFRNISGKQSSLPAMSKVRRLLHYEGLIFRFKFLMLITLA
CAAMTVIFFIVSQVTEGHWKKGWGVTVQVNSAFFTGIIYGMWNLYVFALMFLYAPSHKNGYGEDQS
NGDLGVHSGEELQLTTTITHVDGPTEIYKLTIRKEAQE

Important features of the protein:**Signal peptide:**

amino acids 1-42

Transmembrane domains:amino acids 239-253, 269-284, 302-318, 338-352, 377-399, 434-452,
471-488**N-glycosylation sites.**

amino acids 8-12, 406-410

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 254-258

N-myristoylation sites.amino acids 223-229, 274-280, 305-311, 358-364, 374-380, 386-392,
509-515

15/550

FIGURE 15

GTGAGGGGAACAGCTGATCCGCTCTGTTGGGAGGACAGATATCTCAAGGCCAGGATGGGAAGAAT
CACCCTAAGCCGGGCACCATCCCGTGGTGGAGTCAACTTTCTCAATGTAGCCCGGACCTACA
TCCCCAACACCAAGGTGGAATGTCACTACACCCTTCCCCAGGCACCATGCCAGTGCCAGTG
ACTGGATTGGCATCTTCAAGGTGGAGGCTGCCTGTGTTTCGGGATTACCACACATTTGTGTGGT
CTTCCGTGCCTGAAAGTACAACCTGATGGTTCCCCATTACACACAGTGTCCAGTTCCAAGCCA
GCTACCTGCCCCAAACCAGGAGCTCAGCTCTACCAGTTCCGATATGTGAACCGCCAGGGCCAGG
TGTGTGGGCAGAGCCCCCTTTCCAGTTCCGAGAGCCAAGGCCCATGGATGAACTGGTGACCC
TGGAGGAGGCTGATGGGGGCTCTGACATCCTGCTGGTTGTCCCAAGGCAACTGTGTTACAGA
ACCAGCTCGATGAGAGCCAGCAAGAACGGAATGACCTGATGCAGCTGAAGCTACAGCTGGAGG
GACAGGTGACAGAGCTGAGGAGCCGAGTGCAGGAGCTCGAGAGGGCTCTGGCAACTGCCAGGC
AGGAGCACACGGAGCTGATGGAACAGTACAAGGGGATTTCCCGGTCCCATGGGGAGATCACAG
AAGAGAGGGACATCCTGAGCCGGCAACAGGGAGACCATGTGGCACGCATCCTGGAGCTAGAGG
ATGACATCCAGACCATCAGTGAGAAAGTGCTGACGAAGGAAGTGGAGCTGGACAGGCTTAGAG
ACACAGTGAAGGCCCTGACTCGGGAACAAGAGAAGCTCCTTGGGCAACTGAAAGAAGTACAAG
CAGACAAGGAGCAAAGTGAGGCTGAGCTCCAAGTGGCACAACAGGAGAACCATCACTTAAAT
TGGACCTGAAGGAGGCCAAGAGCTGGCAAGAGGAGCAGAGTGCTCAGGCTCAGCGACTGAAAG
ACAAGGTGGCCAGATGAAGGACACCCTAGGCCAGGCCAGCAGCGGGTGGCCGAGCTGGAGC
CCTTGAAGGAGCAGCTTCGAGGGGCCAGGAGCTTGCAGCCTCAAGCCAGCAGAAAGCCACCC
TTCTTGGGGAGGAGTTGGCCAGTGCAGCAGCAGCCAGGGACCGCACCATAGCCGAACTACACC
GCAGCCGCCTGGAAGTGGCTGAAGTTAACGGCAGGCTGGCTGAGCTCGGTTTGCACCTGAAGG
AAGAAAAATGCCAATGGAGCAAGGAGCGGGCAGGGCTGCTGCAGAGTGTGGAGGCAGAGAAGG
ACAAGATCCTGAAGCTGAGTGCAGAGATACTTCGATTGGAGAAGGCAGTTCAGGAGGAGAGGA
CCCAAACCAAGTGTTCAAGACTGAGCTGGCCCGGGAGAAGGATTCTAGCCTGGTACAGTTGT
CAGAAAGTAAGCGGGAGCTGACAGAGCTGCGGTGAGCCCTGCGTGTGCTCCAGAAGGAAAAGG
AGCAGTTACAGGAGGAGAAACAGGAATTGCTAGAGTACATGAGAAAGCTAGAGGCCCGCCTGG
AGAAGGTGGCAGATGAGAAGTGAATGAGGATGCCACCACAGAGGATGAGGAGGCCGCTGTGG
GGCTGAGCTGCCCCGCAGCTCTGACAGACTCAGAGGACGAGTCCCAGAAGACATGAGGCTCC
CACCTATGGCCTTTGTGAGCGTGGAGACCCAGGCTCCTCTCCTGCTGGGCCTCGAGAGGCTT
CTCCCCCTTGTTGTCATCAGCCAGCCGGCTCCCATTTCTCCTCACCTCTCTGGGCCAGCTGAGG
ACAGTAGCTCTGACTCGGAGGCTGAAGATGAGAAGTCACTCCTGATGGCAGCTGTGCAGAGTG
GGGGTGAGGAGGCCAACTTACTGCTTCCTGAACTGGGCAGTGCCTTCTATGACATGGCCAGTG
GCTTTACAGTGGGTACCCTGTCAGAAACCAGCACTGGGGGCCCTGCCACCCCCACATGGAAGG
AGTGTCCTATCTGTAAGGAGCGCTTTCCTGCTGAGAGTGACAAGGATGCCCTGGAGGACCACA
TGGATGGACACTTCTTTTTTCAGCACCCAGGACCCCTTCACCTTTGAGTGATCTTACTCCCTCG
TACATGCACAAATACACACTCATGCACACACACACTCACACACATGCATACACTTAGGTTTCA
TGCCCATTTTCTATCACACTGGGCTCCATGATATTCTGTTCCCTAAGAACTGCTTCTGTGTGC
CCTGTTTTCTATCCCAAGATTTCTCACTTCATCCTCTCCTACCTGGCTCTTTTGTCCCAGGGAG
GGGTCTGTTCGGAAGCAGTGGCTGAATTTATCCCTGAAAGTGGTTTTGGAGGAACCGGGAT
GGAGGAGGCCCTTCCCCTGTGGGAATAGAATCGTCCACTCCTAGCCCTGGTTGCTTCTGATACA
CAGCCACTGCACACACACACTCACACTCACACTCCCTTGTCTGATGCCCCAAAGCCAATTCTCT
GGGGCACCCCTACCCTCTCTTATTTGAGTTTCCGTGGTTTACCTGAGTTTTCTCTGGGGCTCT
GCACAGAGGCTCAGCATGGACATCATGGCCCTCTCAGGTCCCTTTTGGTTCTCAGTTTCATTG
GTTCCCTCTTTCTGTTCCTCCCATTTGACTTCTGTGCCCCACCCTAGCCTTTTCCATAACCTTAGG
TATTAGTTTTGGAGGGGTTTTTTGTATTTTTGAGGATTCCTGTATTCTGTATCCTCTCCTCGC
ATCTCCTCACATGGAAAGAAATAATGTATTTGTGCCTTCTGTGAGGAATGGGGGGAACAAGTG
GTCCCAGGTATCCCCATTTCCAAGGCCCCCTCCCTCTCCAGGTCCCCCACAGCAATAAAAG
CTTCCCCCTGATATCCATCCCTTTGTAGTTTGAACAAATATATTTATATGATATGTAA

16/550

FIGURE 16

MEESPLSRAPSRGGVNFNLNVARTYIPNTKVECHYTLPPGTMPASDWDWIGIFKVEAACVRDYHT
FVWSSVPESTTDGSPIHSTSVQFQASYLPKPGAQLYQFRYVNRQGQVCGQSPPFQFREPRPMDE
LVTLEEADGGSDILLVVPKATVTLQNQLDESQQERNLMLQLQLLEGQVTELRSRVQELERALA
TARQEHTELMEQYKGISRSHGEITEERDILSRQQGDHVARILELEDDIQTISEKVLTKVELD
RLRDTVKALTREQEKLGLQLKEVQADKEQSEAEQVAQQENHHLNLDLKEAKSWQEEQSAQAQ
RLKDKVAQMKDTLGQAQQRVAELEPLKEQLRGAQELAASSQQKATLLGEELASAAAARDRTIA
ELHRSRLEVAEVNGRLAELGLHLKEEKQOWSKERAGLLQSVEAEKDKILKLSAEILRLEKAVQ
EERTQNQVFKTELAREKDSSSLVQLSESKRELTELRSALRVLQKEKEQLQEEKQELLEMYMRKLE
ARLEKVADEKWNEDATTEDEEAAVGLSCPAALTDSEDESPEDMRLPPYGLCERGDGPGSSPAGP
REASPLVVISQPAPISPHLSGPAEDSSSDSEAEDKSVLMAAVQSGGEEANLLLPELGSAFYD
MASGFTVGTLSETSTGGPATPTWKECPICKERFPAESDKDALEDHMDGHFFFSTQDPFTFE

Important features:**Casein kinase II phosphorylation sites:**

amino acids 28-31, 43-46, 68-71, 72-75, 129-132, 156-159, 208-
211, 239-242, 282-285, 305-308, 376-379, 383-383, 468-471, 520-
523, 521-524, 537-540, 539-542, 543-546, 593-596, 595-598, 597-
600, 612-615, 639-642, 652-655, 667-670, 683-686

N-myristoylation sites:

amino acids 39-44, 107-112, 204-209, 414-419, 561-566, 613-618

Cell attachment sequence:

amino acids 557-559

Leucine zipper pattern sequence:

amino acids 163-184, 475-496, 482-503

17/550

FIGURE 17

GCAAGTTGGGAATTTTAGACTGTCACTGCACATGGACCTCTGGGAAGACGTCTGGCGAGAGCT
AGGCCCACTGGCCCTACAGACGGATCTTGCTGGCTCACCTGTCCCTGTGGAGGTTCCCCTGGG
AAGGCAAGATGCCCAACAACAGCACTGCTCTGTCATTGGCCAATGTTACCTACATCACCATGG
AAATTTTCATTGGACTCTGCGCCATAGTGGGCAACGTGCTGGTCATCTGCGTGGTCAAGCTGA
ACCCCAAGCCTGCAGACCAACACCTTCTATTTTCATTGTCTCTCTAGCCCTGGCTGACATTGCTG
TTGGGGTGCTGGTCATGCCTTTGGCCATTGTTGTCAGCCTGGGCATCACAATCCACTTCTACA
GCTGCCTTTTTATGACTTGCCCTACTGCTTATCTTTACCCACGCCTCCATCATGTCTTGCTGG
CCATCGCTGTGGACCGATACTTGCGGGTCAAGCTTACCGTCAGATTCAGAATTCCTGGGCTCC
CTGGGTGCATTCTATCATTCCAGTTGAAAGTTTGCTTCCTTCCAGTCATGTGGCTCTTCATTC
TACTCTCCTTGCTCTCATTTTCAGATGCCATGGTCATGGATGAAAAGGTCAAGAGAAGCTTTG
TGCTGGACACGGCTTCTGCCATCTGCAACTACAATGCCCACTACAAGAATCACCCCAAATACT
GGTGGCCGAGGCTATTTCCGTGACTACTGCAACATCATCGCCTTCTCCCCTAACAGCACCAATC
ATGTGGCCCTGAGGGACACAGGGAACCAGCTCATTGTCACTATGTCCTGCCTGACCAAAGAGG
ACACGGGCTGGTACTGGTGTGGCATCCAGCGGGACTTTGCCAGGGATGACATGGATTTTACAG
AGCTGATTGTAAGTACGACAAAGGAACCCCTGGCCAATGACTTTTGGTCTGGGAAAGACCTAT
CAGGCAACAAAACCAGAAGCTGCAAGGCTCCCAAAGTTGTCCGCAAGGCTGACCGCTCCAGGA
CGTCCATTCTCATCATTTGCATACTGATCACGGGTTTGGGAATCATCTCTGTAATCAGTCATT
TGACCAAAAGGAGGAGAAGTCAAAGGAATAGAAGGGTAGGCAACACTTTGAAGCCCTTCTCGC
GTGTCCTGACTCCAAAGGAAATGGCTCCTACTGAACAGATGTGACTGAAGATTTTTTTAATTT
AGTTCATAAAGTGATGCTACAACAGAATAATCACCATGACAACTGGCCCACACCTCAGAGACT
GATTCTGATCTCCAGGAATTCTGAAGGACCTCTATCCTTGACAACAATCATTGTCAGCCAG
GTAGCAACGGCGGTAGTCAGAGGAGCTATGATAGACCACACCCAAGCAAGGCTGCCCTCAAAT
AACATCTCAAGATCTTAGTTCTTATGCATTCCATCAGTCAGAAGTGAAGAAGAGGTGGAGAAT
CTGGATTGGGGACCAGGAAATCACTTGTATTTTGTAGCCAATAAATTCCTAGCCAGTGTTGA
ATGAAAAAAAAAAAAA

18/550

FIGURE 18

MPNNSTALSLANVTYITMEIFIGLCAIVGNVLVICVVKLNPSLQTTTFYFIVSLALADIAVGV
LVMP LAIVVSLGITIHFYSCLFMTCLLLIFTHASIMSL LAIVDRYLRVKLTVRFRIPGLPGC
ILSFQLKVCFLPVMWLFILLSLALISDAMVMDEKVKRSFVLDTASAICNNAHYKNHPKYWCR
GYFRDYCNIIAFSPNSTNHVALRDTGNQLIVTMSCLTKEDTGWYWCGIQRDFARDDMDFTELI
VTDDKGTLANDFWSGKDLGKTRSKAPKVVRKADRSRTSILIICILITGLGIISVISHLTK
RRRSQRNRRVGNLTKPF SRVLTPKEMAPTEQM

Important features of the protein:**Transmembrane domains:**

amino acids 16-35, 62-80, 89-101, 134-152, 292-311

N-glycosylation sites.

amino acids 3-7, 4-8, 12-16, 204-208, 273-277

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 316-320

N-myristoylation sites.

amino acids 122-128, 125-131, 258-264

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 214-225

G-protein coupled receptors proteins.

amino acids 29-59, 76-116

19/550

FIGURE 19

CTCGGGCGCGCACAGGACAGTCGGTTTGCCTTGCATTGAGCTGCGGGTTCGGGCCGGCGCCGCTCTCCAATG
GCAATGTGTGTGGCTGGAGGCGAGCGAGGCTTTTCGGCAAAGGCAGTCGAGTGTTCGAGACGGGGCGAGTC
CTGTGAAAGCAGATAAAAGAAAACATTTATTAACGTGTCATTACGAGGGGAGCGCCCGGGCGGGCTGTGCGACT
CCCCGCGGAACATTTGGCTCCCTCCAGCTCCGAGAGAGGAGAAGAAGAAAGCGGAAAAGAGGCAGATTCACGTCTG
TTCCAGCCAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCACGATATTTGATTTATTAGCGATGCCCCCTG
GTTTGTGTGTTACGCACACACACGTGCACACAAGGCTCTGGCTCGCTTCCCTCCCTCGTTTCCAGCTCCTGGGCG
AATCCACATCTGTTTCAACTCTCCGCCGAGGGCGAGCAGGAGCGAGAGTGTGTCGAATCTGCGAGTGAAGAGGG
ACGAGGGAAAAGAAACAAAGCCACAGACGCAACTTGAGACTCCCGCATCCCAAAAGAAGCACCAGATCAGCAAAA
AAAGAAGATGGGCCCCCGAGCCTCGTGCTGTGCTTGTGTCGCAACTGTGTTCTCCCTGTGGGTGGAAGCTC
GGCCTTCCCTGTGCGACCAACCGCCTGAAAGGCAGGTTTCAGAGGGACCGCAGGAACATCCGCCCAACATCATCCT
GGTGCTGACGGACGACAGGATGTGGAGCTGGGTTCATGCAAGTGAACAAGACCGGGCGCATCATGGAGCA
GGCGGGGGCGCACTTCATCAACGCCTTCGTGACCACACCCATGTGCTGCCCTCAGCTCCTCCATCCTCACTGG
CAAGTACGTCCACAACCACAACACCTACACCAACAATGAGAACTGCTCCTCGCCCTCCTGGCAGGCACAGCACGA
GAGCCGCACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTTAATGAATA
CAACGGCTCCTACGTGCCACCGGGTGAAGGAGTGGGTGCGACTCCTTAAAACTCCCGCTTTTATAACTACAC
GCTGTGTGGAACGGGGTGAAGAGAAGCAGGCTCCGACTACTCCAAGGATTACCTCACAGACCTCATACCAA
TGACAGCGTGAGCTTCTTCGCGACGTCCAAGAAGATGTACCGCACAGGCCAGTCTCATGGTCACTAGCCATCG
AGCCCCCACGGCCCTGAGGATTAGCCCCACAATATTACGCCTCTTCCCAAACGCATCTCAGCACATCAGGCC
GAGCTACAACCTACGCGCCCAACCCGACAAACACTGGATCATGCGCTACACGGGGCCCATGAAGCCCATCCACAT
GGAATTCACCAACATGCTCCAGCGGAAGCGCTTGACAGCCCTCATGTGCGTGGACGACTCCATGGAGACGATTTA
CAACATGCTGGTTGAGACGGGCGAGCTGGACAACACGTACATCGTATACACCGCGGACACGGTTACCACATCGG
CCAGTTTGGCCTGGTGAAAGGGAAATCCATGCCATATGAGTTTGACATCAGGGTCCCGTTTACGTGAGGGGCC
CAACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCTCAACATTGACCTGGCCCCACCATCCTGGACATTGC
AGGCCTGGACATACCTGCGGATATGGACGGGAAATCCATCCTCAAGCTGCTGGACACGGAGCGGCCGGTGAATCG
GTTTCACTTGAAAAAGAAGATGAGGCTGTGGCGGACTCCTTCTTGGTGAGAGAGGCAAGCTGCTACACAAGAG
AGACAATGACAAGGTGGACGCGCCAGGAGGAGAATTTCTGCCAAGTACCAGCGTGTGAAGGACCTGTGTACGG
TGCTGAGTACCCAGCGCGCTGTGACGAGCTGGGACAGAAGTGCCAGTGTGTGGAGGACGCCACGGGAAGCTGAA
GCTGCATAAGTGCAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCCTCTCCAACCTCGTGCCCAAGTACTACGG
GCAGGGCAGCGAGGCTGCACCTGTGACAGCGGGGACTACAAGCTCAGCCTGGCCGACGCCGAAAAAACTCTT
CAAGAAGAAGTACAAGGCCAGCTATGTCCGCACTGCTCCATCCGCTCAGTGGCCATCGAGGTGGACGGCAGGGT
GTACCAGTAGGCTGGGTGATGCCGCCAGCCCCGAAACCTCACCAAGCGGCACTGGCCAGGGGCCCTGAGGA
CCAAGATGACAAGGATGTTGGGGACTTCAGTGGCACTGGAGGCCTTCCGACTACTCAGCCGCCAACCCCATTA
AGTGACACATCGGTGCTACATCCTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTGCAGGC
CTGGAAAGACCACAAGCTGCACATCGACCACGAGATTGAAACCTGCAGAACAAAATTAAGAACCTGAGGGAAAT
CCGAGGTCACTGAAGAAAAAGCGGCCAGAAGATGTGACTGTACAAAATCAGCTACCACACCCAGCACAAAGG
CCGCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTAGGAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTT
CGGGAGCAGAAGCGCAAGAAGAACTCCGCAAGCTGCTCAAGCGCTGCAGAACAAACGACACGTGCAGCATGCC
AGGCCTCAGTGCTTACCCACGACAACCAGCACTGGCAGACGGCGCCTTTCTGGCACTGGGGCCTTTCTGTGC
CTGCACCAGCGCCAACAATAACAGTACTGGTGATGAGGACCATCAATGAGACTCACAATTTCTCTTCTGTGA
ATTTGCAACTGGCTTCTAGAGTACTTTGATCTCAACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACT
GGACAGGGATGTCTCAACAGCTACACGTACAGCTCATGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAA
CCCCGACTCGAAACATGGACCTGGATGGAGGAAGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCC
AGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACGTGGGAAGGCTGGGAAGGTTAAGAAACAACAGAGGT
GGACCTCCAAAAACATAGAGGCATCACCTGACTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACC
TGTGCTATTGGCCAGGAGGCTGAGAAAGCAAGCAGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACCA
GCAGGAGCAGAGATACTTCAGGAAGTCCATTTTGGCCCTGCTTTTGCTTTGGATTATACCTCACCAGCTGCAC
AAAATGCATTTTTTCTGATCAAAAAGTCACCACTAACCTCCCCAGAAGCTCACAAAGGAAAACGGAGAGAGCG
AGCGAGAGAGATTTCTTGGAAATTTCTCCCAAGGGCGAAAGTCAATGGAATTTTAAATCATAGGGGAAAAGCA
GTCCTGTTCTAAATCCTCTTATTCTTTGGTTTGTACAAAGAAGGAACTAAGAAGCAGGACAGAGGCAACGTGG
AGAGGCTGAAAACAGTGCAGAGACGTTTGACAAATGAGTCAGTAGCACAAAAGAGATGACATTTACCTAGCACTAT
AAACCTGGTTGCCTCTGAAGAACTGCCTTCATTGTATATATGTGACTATTTACATGTAATCAACATGGGAACT
TTTAGGGGAACCTAATAAGAAATCCCAATTTTCAGGAGTGGTGGTGTCAATAACGCTCTGTGGCAGTGTAAAA
GAAAA

20/550

FIGURE 20

MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQRDRRNIRPNIILVLTDDQDVELGSMQV
MNKTRRIMEQGGAHFINAFVTTMCCPSRSSILTGYVHNHNTYTNNENCSSPSWQAQHESTR
FAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHGSDYSK
DYLTDLITNDSVSFFRTSKKMPHRPVL MVISHAAPHGPEDSAPQYSRLFPNASQHITPSYNY
APNPDKHWMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNMLVETGELDNTYIVYT
ADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPVNEAGCLNPHIVLNIDLAPTILDIAGLDIP
ADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEENFLPKYQR
VKDLCQRAEYQTACEQLGQKWQCVEDATGKLLKHCKGPMRLGGSRALSNLVPKYYGQGSEAC
TCDSGDYKLSLAGRRKKLFKKKYKASYVRSRSIRSAIEVDGRVYHVGLGDAAQPRNLTKRHW
PGAPEDQDDKDGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKSLQAWKDHKLH
IDHEIETLQNKIKNLREVRGHLKKRPEECDCHKISYHTQHKGRLKHRGSSLHPFRKGLQEKD
KVWLLREQKRKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLGPFCACTSANNT
YWCMTINETHNFLCFEATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGY
KQCNPRTRNMDLDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG

Important features:**Signal peptide:**

amino acids 1-17

Sulfatases signature 1.

amino acids 86-99

Homologous region to sulfatase:

amino acids 87-106, 133-146, 216-229, 291-320, 365-375

N-glycosylation sites.amino acids 65-69, 112-116, 132-136, 149-153, 171-175, 198-202,
241-245, 561-565, 608-612, 717-721, 754-758, 764-768

21/550

FIGURE 21

GGGCGCGGAGAGCTGCTAGGGCGGTTTCTCTGCCTCGGGCCTGTTGGGCAGGGCCGGCT
AAGGTGCGCGTGCTCGCTGGTTCTAACCCCTTCTGTTGGGCGTTTCTGCTGAGAGGCGGGA
GGCGCTGAGAGTCTGTGCGGAGGTCCGTGGACAGACTGCTTTGCTCGTTGTTGCTCTTCG
GAGGCGGCGATCCCCGAAGGCGAGCTGAAATACGGCTGCAGGCTACAATTTGCAGCCGAC
GATTATGGAAGACGGAAGCGGGAGAGGTGGCCACCCCTCATGGAGCGCTTGTGCTCGGAT
GGCTTCGCATTTCCCCAATACCCCATTAACCGTATCATCTGAAGAGGATCCACAGAGCT
GTCTTACATGGTAATCTAGAGAACTGAAGTACCTTCTGCTCACGTATTATGACGCCAAT
AAGAGAGACAGGAAGGAAAGGACCGCCCTACATTTGGCCTGTGCCACTGGCCAACCGGAA
ATGGTACATCTCCTGGTGTCCAGAAGATGTGAGCTTAACCTCTGCGACCGTGAAGACAGG
ACACCTCTGATCAAGGCTGTACAACCTGAGGCAGGAGGCTTGTGCAACTCTTCTGCTGCAA
AATGGCGCCAATCCAAATATTACGGATTTCTTTGGAAGGACTGCTCTGCACTACGCTGTG
TATAATGAAGATACATCCATGATAGAAAACTTCTTTCACATGGTACAAATATTGAAGAA
TGCAGCAAGGTATTAGGTCAACCAATGTTATTTTCAAACCTATCTGAAATGAATTTATTTTA
ACATTGACACATGTAAGGGTCAATTTTTTCATATTTGGAAGCTCAAACATTCCTTGAATGA
AAATATTTTGAAATGCCTTAACTGTCTAAGATTTTACTTTAAATATTGGAACTTTAAAG
AAGCATTATAGGGAACAGCCTTTTTTCATGCACTTATGGTAAATAACTATAAAAACAAAT
GAATTACAATAAATTTATAATTCATGACAACCTGAATTTGGGAAAGGTAATAGTTAAGTGT
TTTTCCACTAAATTACTTTTT

22/550

FIGURE 22

MERLCSDGFAFPQYPIKPYHLKRIHRAVLHGNLEKLKYLLTTYDANKRDRKERTALHLACAT
GQPEMVHLLVSRRCELNLCDREDRTPLIKAVQLRQEACATLLLQNGANPNITDFFGRTALHYA
VYNEDTSMIEKLLSHGTNIEECSKV

Important features of the protein:

N-glycosylation site.

amino acids 113-117

N-myristoylation site.

amino acids 109-115

Microbodies C-terminal targeting signal.

amino acids 149-153

23/550

FIGURE 23

GAGGCAGAAAGGCAGAAAGGAGAAAATTCAGGATAACTCTCCTGAGGGGTGAGCCAAGCCCTG
CCATGTAGTGCACGCAGGACATCAACAAACACAGATAACAGGAAATGATCCATTCCCTGTGGT
CACTTATTCTAAAGGCCCAACCTTCAAAGTTCAAGTAGTGATATGGATGACTCCACAGAAAG
GGAGCAGTCACGCCTTACTTCTTGCCCTAAGAAAAGAGAAGAAATGAACTGAAGGAGTGTGT
TTCCATCCTCCCACGGAAGGAAAGCCCCCTCTGTCCGATCCTCCAAAGACGGAAAGCTGCTGGC
TGCAACCTTGCTGCTGGCACTGCTGTCTTGCTGCCTCACGGTGGTGTCTTTCTACCAGGTGGC
CGCCCTGCAAGGGGACCTGGCCAGCCTCCGGGCAGAGCTGCAGGGCCACCACGCGGAGAAGCT
GCCAGCAGGAGCAGGAGCCCCCAAGGCCGGCCTGGAGGAAGCTCCAGCTGTCACCGCGGGACT
GAAAATCTTTGAACCACCAGCTCCAGGAGAAGGCAACTCCAGTCAGAACAGCAGAAATAAGCG
TGCCGTTCAAGGTCCAGAAGAAACAGTCAAGACTGCTTGCAACTGATTGCAGACAGTGA
AACACCAACTATACAAAAAGGATCTTACACATTTGTTCCATGGCTTCTCAGCTTTAAAGGGG
AAGTGCCCTAGAAGAAAAAGAGAATAAAATATTGGTCAAAGAACTGGTTACTTTTTTATATA
TGGTCAGGTTTTATATACTGATAAGACCTACGCCATGGGACATCTAATTCAGAGGAAGAAGGT
CCATGTCTTTGGGGATGAATTGAGTCTGGTGACTTTGTTTCGATGTATTCAAAATATGCCTGA
AACACTACCCAATAATTCTTGCTATTTCAGCTGGCATTGCAAACTGGAAGAAGGAGATGAACT
CCAATTGCAATACCAAGAGAAAATGCACAAATATCACTGGATGGAGATGTCACATTTTTTGG
TGCATTGAACTGCTGTGACCCTACTTACACCATGTCTGTAGCTATTTTCCTCCCTTTCTCTGT
ACCTCTAAGAAGAAAGAATCTAACTGAAAATACCAAAAAAAAAAAAAAAAAA

24/550

FIGURE 24

MDDSTEREQSRLTSCCLKKREEMKLKECVSILPRKESPSVRSSKDGKLLAATLLLALLSCCLTV
VSFYQVAALQGD LASLRAELQGHHA EKLPAGAGAPKAGLEEAPAVTAGLKIFEPPAPGEGNSS
QNSRNKRAVQGP EETVTQDCLQLIADSETPTIQKGSYTFVPWLLSFKRGSAL EEEKENKILVKE
TGYFFIYGQVLYTDKTYAMGH LIQRKKVHVFGDELSLVTLFRCIQNMPETLPNNSCYSAGIAK
LEEGDELQLAIPRENAQISLDGDVTFFGALKLL

Transmembrane domain:

amino acids 47-72

N-glycosylation site.

amino acids 124-127, 242-245

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 33-36, 173-176

N-myristoylation site.

amino acids 96-101

TNF family proteins.

amino acids 172-206

25/550

FIGURE 25

CTGCTTGGATACCTCCAGTCCCCAACTGTGTTCCAGGAGTTTTCTTGGCCGAAGCTGCCCCGA
TGTTTGGAGCCTTTTCTTCCCAGAGAAGAAGATGGACTGAAAGCTGCCAGTTGGGGACTTTTTTG
TGATCACGGCGTTGCAGCGTTTTAAAGGAGGTGATGGGGCTTGCCTGGCTTGTCTTCCCACC
CAAGTGAAGAGTTGATGTTCACTGGTTATGCTTAGACAATGTGCAGTTTGTGTTAATTTAAAA
TTTTGGGTGGGATAGGGGCATAGGCTTGTGAAGGGCAGTCCGGATCCGGAGGAACCTCGTCTTT
GTCCCTGGTAGGAGAGACACCCCACTATCCTCGATGCCGTCAGCCTTGGCCATCTTCACT
TGCCGCCCCGAACCTCGCACCCGTTTCAGGAGCGTCATGTCTACCTGGACGAGCCCATCAAATC
GGCCGCTCAGTGGCCCGCTGTGACACAGCGCAGAATAATGCCACTTTTGATTGCAAAGTGCTA
TCAAGGAACCACGCTCTCGTCTGGTTTGATCACAAGACGGGCAAGTTTTATCTTCAAGACACT
AAAAGTAGTAATGGTACTTTTATAAATAGCCAGAGATTGAGTCGAGGCTCTGAAGAAAGTCCA
CCATGTGAAATTCCTTTCCGGTGACATTATCCAGTTTGGAGTAGACGTGACAGAGAATACACGG
AAAGTTACCCATGGGTGTATTGTTTCCACAATAAACTTTTTCTACCAGATGGTATGGAAGCC
CGGCTCCGCTCAGATGTCATCCATGCACCAATTACCAAGTCCTGTTGACAAAGTTGCTGCTAAC
ACTCCAAGTATGTACTCTCAGGAACATAATCCAGCTTTCTCAGTATCTACAGGAGGCCTTACAT
CGGGAACAAATGTTGGAACAGAAGTTAGCCACGCTTCAGCGGCTACTAGCCATCACCCAAGAG
GCTTCAGATACCAGTTGGCAGGCTTTAATAGATGAAGATAGACTCTTATCACGGTTAGAAGTT
ATGGGAAACCAATTACAGGCATGCTCCAAAAATCAAACAGAAGATAGTTTACGAAAGGAACCTT
ATAGCATTACAAGAGGATAAACATAACTATGAGACAACAGCCAAAGAGTCCCTGAGGCGGGTT
CTTCAGGAGAAAATTGAAGTGGTTAGAAAACCTTTCAGAAGTTGAGCGAAGTCTGAGTAATACT
GAAGATGAATGTACCCATCTGAAAGAAATGAATGAAAGGACTCAGGAAGAATTAAGAGAATTA
GCCAACAAATATAATGGAGCAGTTAATGAGATTAAAGATTTATCTGATAAATTAAGGTAGCA
GAGGGAAAACAAGAGGAAATCCAACAGAAGGGACAGGCTGAGAAAAAAGAATTACAACATAAA
ATAGATGAAATGGAAGAAAAAGAACAGGAGCTCCAGGCAAAAATAGAAGCTTTGCAAGCTGAT
AATGATTTACCAATGAAAGGCTAACAGCTTTACAAGTACGGTTAGAACATCTTCAGGAGAAA
ACTCTTAAAGAATGCAGCAGCTTGGCTGATCGTCGAAGGGCATCTAACCAAGCGGTAGAAGA
AACAAAGCTTTCAAAGGTTTGTCTTCTGTTTTCTATGTTTTTTGACAGTTCTTTTGGATAA
TGAAGGTTAGTGATATATTTTCAAGGTTATAGTATTTTAAACCATCAGTTTACTTCTTATAGCTC
ACAAAATAGCAAGCCAGTAACAGTATCAGATAATATATAAAATAATCAGACTTCTGTTTTAAG
AAGGGTATCGTAACCTGGAATGTGTCTTTTTAAGTGGATGTATATTTATGGTTTTTTGAATGTT
AGTACTTGATATAGGTTTCTTTAGGTATTAAAGATTTGTTGCAATCTCTGTCAATCCCAGCAT
TAATTTTCAGCTTTGATCTCAAATTTTAAATCAAACACAATGTAAGTCGTTTGTGATACAACCTTA
AGTGAAACATGCTTGCACCTTCTATTTTGGGGGTACAGTACCTTTAAATCTCTTATGATGTT
TAATATTTCTTAATTTTGGCATCTCAGTTTGATTTAAACAAAATTAATGACTTTTGTGAAT
GTAGAATCTTCTTATATTTTATGAGTAGTCCAGTAATTGCCCAAAGTAGTTTATTGTGTTAAT
TCTGTTACAGTTGTCAGAGAAGAAAAGTGAGTTTTAAAGCACCATATTGTCAAGTCACTTTTA
TACATAGGGAAATTAGGCAAATAAATTTGGTGGCATGTGTTTATCATAGTAGAACTTTCATTA
GACTATACCAGTATAAAATTTAAACTAGATTACAGTCCTTTTGGCCAATTAACATTGAG
TTACAAAAGTTTGAGATACTTAATTTTAGTACATTCTATTTTATTAAAGTAACTGGATTCAAT
TGACTTTTTTAAACCATGTAAGAGGATGGTGTTATTTCAAATATCTCGTGGTTTTCCATTCTGAA
TTTTGTGCACGGCAGATGCCATATTTGGGGAAAAATGCATAGAATATGCATCATTAATATTG
TTTTGGCAAACAGGCATTGAGTTTCAGAACAGTGAACATAATTTTAGTACATATGGCAATTTTT
TTTACCCTTATTAAAGTGAGATGAGAACAGACCTTAAATAGCTTTTACCTCACCATCCAAATA
CCTATTCAGATTAGTTGGTTGAATAGCCAGCACTTTGAAGTAGAGCCTTAGG

26/550

FIGURE 26

MEARLRSDVIHAPLPSPVDKVAANTPSMYSQELFQLSQYLQEALHREQMLEQKLATLQRLAI
TQEASDTSWQALIDEDRLLSRLEVGMGNQLQACSKNQTEDSLRKELIALQEDKHNHYETTAKESL
RRVLQEKIEVVRKLSEVERSLSNTEDECTHLKEMNERTQEELRELANKYNGAVNEIKDLSDKL
KVAEGKQEEIQQKGQAEKKELQHKIDEMEEKEQELQAKIEALQADNDFTNERLTALQVRLEHL
QEKTLECCSSLADRRRRASNQSGRRNKAFKRFFVFCFSMFFDSSFG

Important features of the protein:**N-glycosylation sites.**

amino acids 98-102, 271-275

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 138-142, 267-271

Amidation site.

amino acids 273-277

Tropomyosins proteins.

amino acids 169-217

27/550

FIGURE 27

GAACCTGGCGCCGCGGAACTGATCGCGGCCTAGTCCCGACGCGTGTGTGCTAGTGAGCCGGA
GCCGGCGACGGCGGCAGTGGCGGCCCGGCCTGCAGGAGCCCGACGGGGTCTCTGCCATGGGGG
AGTGACGCGCCTGCACCCGCTGTTCCGCGGCAGCGGCGAGACATGAGGAGACCCCGCGACAGG
GGCAGCGGGCGGCGGCTCGTGAGCCCCGGGATGGAGGAGAAATACGGCGGGGACGTGCTGGCCG
GCCCCGGCGGCGGCGGCGGCCTTGGGCCGGTGACGTACCCAGCGCTCGATTAACAAAATATA
TTGTGTTACTATGTTTCACTAAATTTTTGAAGGCTGTGGGACTTTTCGAATCATATGATCTCC
TAAAAGCTGTTACATTGTTTCAGTTTCATTTTTATATTAACAACTGGGACTGCATTTTTATGG
TTTTGTTTCAAAGCCATTTTCTTCTGGGAAAACATTACCAAACACCAGTGGATCAAATAT
TTAAACATGCAGTTGCTGGGTGTATTATTTCACTCTTGTGGTTTTTGGCCTCACTCTTTGTG
GACCACTAAGGACTTTGCTGCTATTTGAGCACAGTGATATTGTTGTCATTTCACTACTCAGTG
TTTTGTTTACCAGTTCTGGAGGAGGACCAGCAAAGACAAGGGGAGCTGCTTTTTTCAATTATTG
CTGTGATCTGTTTATTGCTTTTTGACAATGATGATCTCATGGCTAAAATGGCTGAACACCCTG
AAGGACATCATGACAGTGCTCTAACTCATATGCTTTACACAGCCATTGCCTTCTTAGGTGTGG
CAGATCACAAGGGTGGAGTATTATTGCTAGTACTGGCTTTGTGTTGTAAAGTTGGTTTTTCATA
CAGCTTCCAGAAAGCTCTCTGTGCGACGTTGGTGAGCTAAACGTCTTCAAGCTTTATCTCATC
TTGTTTCTGTGCTTCTCTTGTGCCATGGGTCATTGTTCTTTCTGTGACAACTGAGAGTAAAG
TGGAGTCTTGGTTTTTCTCTCATTTATGCCTTTTTGCAACGGTTATCTTTTTTGTGATGATCCTGG
ATTTCTACGTGGATTCCATTTGTTTCAGTCAAAATGGAAGTTTCCAAATGTGCTCGTTATGGAT
CCTTTCCCATTTTTATTAGTGCTCTCCTTTTTTGGAATTTTTGGACACATCCAATAACAGACC
AGCTTCGGGCTATGAACAAAGCAGCACACCAGGAGAGCACTGAACACGTCCTGTCTGGAGGAG
TGGTAGTGAGTGCTATATTCTTCATTTTGTCTGCCAATATCTTATCATCTCCCTCTAAGAGAG
GACAAAAGGTACCCTTATTGGATATTCTCCTGAAGGAACACCTCTTTATAACTTCATGGGTG
ATGCTTTTTCAGCATAGCTCTCAATCGATCCCTAGGTTTATTAAGGAATCACTAAAACAAATTC
TTGAGGAGAGTGACTCTAGGCAGATCTTTTACTTCTTGTGCTTGAATCTGCTTTTTTACCTTTG
TGGAATTATTCTATGGCGTGCTGACCAATAGTCTGGGCCTGATCTCGGATGGATTCCACATGC
TTTTTGACTGCTCTGCTTTAGTCATGGGACTTTTTGCTGCCCTGATGAGTAGGTGGAAAGCCA
CTCGGATTTTCTCCTATGGGTACGGCCGAATAGAAATTCTGTCTGGATTTATTAATGGACTTT
TTCTAATAGTAATAGCGTTTTTGTGTTTATGGAGTCAGTGGCTAGATTGATTGATCCTCCAG
AATTAGACACTCACATGTTAACACCAGTCTCAGTTGGAGGGCTGATAGTAAACCTTATTGGTA
TCTGTGCCTTTAGCCATGCCCATAGCCATGCCCATGGAGCTTCTCAAGGAAGCTGTCACTCAT
CTGATCACAGCCATTACACCATATGCATGGACACAGTGACCATGGGCATGGTCACAGCCACG
GATCTGCGGGTGGAGGCATGAATGCTAACATGAGGGGTGTATTTCTACATGTTTTGGCAGATA
CACTTGGCAGCATTGGGTGTGATCGTATCCACAGTTCTTATAGAGCAGTTTGGATGGTTCATCG
CTGACCCACTCTGTTCTCTTTCTACTGCTATATTAATATTTCTCAGTGTTGTTCCACTGATTA
AAGATGCCTGCCAGGTTCTACTCCTGAGATTGCCACCAGAATATGAAAAGAACTACATATTG
CTTTAGAAAAGATACAGAAAATTGAAGGATTAATATCATACCGAGACCCTCATTTTTTGGCGTC
ATTCTGCTAGTATTGTGGCAGGAACAATTCATATACAGGTGACATCTGATGTGCTAGAACAAA
GAATAGTACAGCAGGTTACAGGAATACTTAAAGATGCTGGAGTAAACAATTTAACAATTCAAG
TGGAAAAGGAGGCATACTTCAACATATGTCTGGCCTAAGTACTGGATTTTCATGATGTTCTGG
CTATGACAAAACAAATGGAATCCATGAAATACTGCAAAGATGGTACTTACATCATGTGAGATA
ACTCAAGAATTACCCCTGGAGAATAAACAATGAAGATTAAATGACTCAGTATTGTAAATATTG
CCAGAAGGATAAAAATTACACATTAAGTGTACAGAAACAGAGTTCCCTACTACTGGATCAAGG
AATCTTTCTTGAAGGAAATTTAAATACAGAATGAAACATTAATGGTAAAAAAA

28/550

FIGURE 28

MEEKYGGDVLAGPGGGGGGLGPVDVPSARLTKYIVLLCFTKFLKAVGLFESYDLLKAVHIVQFI
FILKLGTAFFMVLFOKPFSSGKTITKHQWIKIFKHAVAGCIISLLWFFGLTLCGPLRTLLLF
HSDIVVISLLSVLFTSSGGGPAKTRGAFFIIAIVICLLLFDDNDDLMAKMAEHPEGHDSALTH
MLYTAIAFLGVADHKGGVLLLVLALCCKVGFTASRKLSVDVGGAKRLQALSHLVSVLLCPW
VIVLSVTTESKVESWFSLIMPFATVIFVVMILDFYVDSICSVKMEVSKCARYGSFPIFISALL
FGNFWTHPITDQLRAMNKAHQESTEHLVSGGVVVSIAIFFILSANILSSPSKRGQKGTIGYS
PEGTPLYNFMGDAFQHSQSIPRFIKESLKQILEESDSRQIFYFLCLNLLFTFVELFYGVLTN
SLGLISDGFHMLFDCSALVMGLFAALMSRWKATRIFSYGYGRIEILSGFINGLFLIVIAFFVF
MESVARLIDPPELDTHMLTPVSVGGLIVNLIGICAFSHAHAHSHAHGASQGSCHSSDHS SHHM
GHSDHGHGSHSGSAGGMNANMRGVFLHVLADTLGSIGVIVSTVLIEQFGWFIADPLCSLSTA
ILIFLSVVPLIKDACQVLLRLPPEYEKELHIALEKIQKIEGLISYRDPHFWRHSASIVAGTI
HIQVTSVDVLEQRIVQQVTGILKDAGVNNLTIQVEKEAYFQHMSGLSTGFHDVLAMTKQMESMK
YCKDGTIIM.

Important features of the protein:**Signal peptide:**

amino acids 1-46

Transmembrane domains:amino acids 59-77, 101-119, 150-167, 205-223, 239-258, 267-284,
305-324, 343-360, 421-440, 452-469, 486-505, 522-539, 592-612,
621-641**N-glycosylation site.**

amino acids 721-725

Glycosaminoglycan attachment site.

amino acids 143-147

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 225-229

Tyrosine kinase phosphorylation sites.

amino acids 750-758, 756-764

N-myristoylation sites.amino acids 14-20, 46-52, 102-108, 112-118, 144-150, 317-323,
347-353, 369-375, 372-378, 437-443, 462-468, 529-535, 549-555,
553-559, 579-585, 582-588, 583-589, 584-590, 605-611, 737-743**Multicopper oxidases protein:**

amino acids 561-569

29/550

FIGURE 29

GGCACGAGGGCAGGATATTAGAAATGGCTACTCCCCAGTCAATTTTCATCTTTGCAATCTGCA
TTTTAATGATAACAGAATTAATTCTGGCCTCAAAAAGCTACTATGATATCTTAGGTGTGCCAA
AATCGGCATCAGAGCGCCAAATCAAGAAGGCCTTTCACAAGTTGGCCATGAAGTACCACCTG
ACAAAAATAAGAGCCCGGATGCTGAAGCAAAATTCAGAGAGATTGCAGAAGCATATGAAACAC
TCTCAGATGCTAATAGACGAAAAGAGTATGATACACTTGGACACAGTGCTTTTACTAGTGGTA
AAGGACAAAGAGGTAGTGGAAAGTTCTTTTGAGCAGTCATTTAACTTCAATTTTGATGACTTAT
TTAAAGACTTTGGCTTTTTTGGTCAAAACCAAAACACTGGATCCAAGAAGCGTTTTGAAAATC
ATTTCCAGACACGCCAGGATGGTGGTTCCAGTAGACAAAGGCATCATTTCCAAGAATTTTCTT
TTGGAGGTGGATTATTTGATGACATGTTTGAAGATATGGAGAAAATGTTTTCTTTTAGTGGTT
TTGACTCTACCAATCAGCATACAGTACAGACTGAAAATAGATTTTCATGGATCTAGCAAGCACT
GCAGGACTGTCACTCAACGAAGAGGAAATATGGTTACTACATACACTGACTGTTTCAGGACAGT
AGTTCTTATTCTATTCTCACTAAATCCAACTGGTTGACTCTTCCTCATTATCTTTGATGCTAA
ACAATTTTCTGTGAACTATTTTGACAAGTGCATGATTTCACTTTAAACAATTTGATATAGCTA
TTAAATATATTTAAGGGTTTTTTTTTTTTTGACAAATTCACATTCAACGAGTAGACAAAATGCT
AATTATTTCCCTGATTAGGAAAGTTTCTTTAAAAACACGTAATTTTGCCTAGTGCTTTTTTCT
CTACCTGCCCTTGGGCTCACTAATATCACCAGTATTATTACCAAGAAAATATTGAGTTTACCT
GATTAACTTTAAAAGTTAATTGTAGATTTAAATTTGTGTGAACCTAATGATTTTTGCAGTGAA
ACCTTTACTAATTCAAAGTTGCATGTTCTATGACATCTGTGACTTGCCTTGCAGAGTGACAT
GAACTGTATAATTGAGTCATTCAGTAAAGGAGAACAGTATCTTGGTTAATTGCTACTGAAAG
GTTGAGAAAGGAATGGTTTGATATTTACCACAGCGCTGTGCCTTTCTACAGTAGAACTGGGGT
AAAGGAAATGGTTTTATTGCCCATAGTCATTTAGGCTGGAAAAAAGTTGAAAACCTAACGAAA
TATTGCCAAGAGATTGTTATGTGTTTGGTTCCAGCCTAAAAATGATTTTGTAGTGTTGAAATC
ATAGCTACTTACATAGCTTTTTTCATATTTCTTTCTTAGTTGTTGGCACTCTTAGGTCTTAGTA
TGGATTTATGTGTTTGTGTGTGTGTAGTTTATCCTCTCTCATCTTTATCTAGAGATTGACT
GATACCTCATTCTGTTTGTAACCAGCCAGTAATTTCTGTGCAACCTTACTATGTGCAATAT
TTTTAAATCCTGAGAAATGTGTGCTTTTGTTCGGATAGACTTATTCTTTAGTTCTGCACT
TTTCCACATTATACTCCATATGAGTATTAATCCTATGGATACATATTAAACAAGTGTCTCAT

30/550

FIGURE 30

MATPQSIFIFAICILMITELILASKSYDILGVPKSASERQIKKAFHKLAMKYHPDKNKSPDA
EAKFREIAEAYETLSDANRRKEYDTLGHSFTSGKGQRGSGSSFEQSFNFNFDDLKDFGFFG
QNQNTGSKKRFEHFQTRQDGGSSRQRHHFQEFSGGGLFDDMFEDMEKMFSGFDSTNQHT
VQTENRFHGSSKHCRVTVTQRRGNMVTITYTDCSGQ

Important features of the protein:**Signal peptide:**

amino acids 1-23

Nt-dnaJ domain signature.

amino acids 27-59, 66-90

Glycosaminoglycan attachment site.

amino acids 96-100

N-myristoylation sites.

amino acids 32-38, 99-105, 102-108, 126-132, 211-217

31/550

FIGURE 31

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTGG
GCAGAAAGGAGGGTGCTTCGGAGCCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACAAT
TCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCTGAG
ATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCAAATG
CAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCTACGCA
TTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCTCAGAACCTCTCTGTA
CTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGAAACAGTG
TACTATTCTGTGCAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCTGGATCCCC
AGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATCACGGCCACT
GTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTGGAGCATCCTG
AAGCATCCCTTTAATAGAACTCAACCATCCTTACCCGACCTGGGATGGAGATCACCAAAGAT
GGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTCCTTGTGGCCTAC
TGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAATGGTGAGGAGTGGGGGTATTCCAGTG
CACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGACATTCGTGAAGGCC
ATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAAGGAGAGGCCATTCCC
CTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGTGGTCGTGCCACTGTTT
GTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCGTGGTGGTCCTCCAGACACC
TTGAAAATAACCAATTACCCCCAGAAGTTAATCAGCTGCAGAAGGGAGGAGGTGGATGCCTGT
GCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGATCTCATAAGTTTTGCGGAAGG
GCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACCATGAGGGGACAAGTTGTGTT
TCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGAGCCTGTTGTCTACAAGTCTAG
AAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACTGACTGAGGCTTAGGGGATGTG
ACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACCCTGGGAAAAGTGACTTCATCCCT
TCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACCTACACACCTGCTAAACACACACAC
ACAGAGTCTCTCTATATATACACACGTACACATAAATACACCCAGCACTTGCAAGGCTAGA
GGGAAACTGGTGACACTCTACAGTCTGACTGATTAGTGTCTTCTGGAGAGCAGGACATAAATG
TATGATGAGAATGATCAAGGACTCTACACACTGGGTGGCTTGGAGAGCCCACTTTCCAGAAT
AATCCTTGAGAGAAAAGGAATCATGGGAGCAATGGTGTGAGTTCACCTCAAGCCCAATGCCG
GTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGTAGGTGACCTGGAGGAAGGTCACAGCCACA
CTGAAAATGGGATGTGCATGAACACGGAGGATCCATGAACTACTGTAAAGTGTGACAGTGTG
TGCACACTGCAGACAGCAGGTGAAATGTATGTGTGCAATGCGACGAGAATGCAGAAGTCAGTA
ACATGTGCATGTTTGTGTGCTCCTTTTTTCTGTTGGTAAAGTACAGAATTCAGCAAATAAAA
AGGGCCACCCTGGCCAAAAGCGGTAAAAA

32/550

FIGURE 32

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLMLWSPVIAPGET
VYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECVDVDDITATVPYNLRVRATLGSQTSAWSI
LKHPFNRNSTILTRPGMEITKDGFLVIELEDLGPQFEFLVAYWRREPGAEHVKMVRSGGIP
VHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPVLALFAFVGFMILVVLVPL
FVWKMGRLLQYSCCPVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:**Signal peptide:**

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation sites.

amino acids 40-44, 134-138

Tissue factor proteins.

amino acids 92-120

Integrins alpha chain proteins.

amino acids 232-263

33/550

FIGURE 33

GAGACACGCGAGCGGGGAGACCTCCAAGGCAGCGAGGCATCGGACATGTGTCAGCACATCTGG
GGCGCACATCCGTCGAGCCCGAGGGGAGATTTGCCGGAACAATTCAAACCTGCGATATTGATCT
TGGGGGTGACTGTCCCTGGCCGGCTGTGCGGTGGGAGTGCGAGTGTGCACTCGCTCGGAAGTG
TGTGCGAGTGTGTATGTGTGTGTGCCGTGTGCGGCTCCCCCTTCCCCCGTTTTCCCGTCGA
GTGATGCACTTGGAATGAGAATCAGAGG**ATG**GAAATAGTCTGGGAGGTGCTTTTTCTTCTTCA
AGCCAATTTTCATCGTCTGCATATCAGCTCAACAGAATTCACCAAAAATCCATGAAGGCTGGTG
GGCATAACAAGGAGGTGGTCCAGGGAAGCTTTGTTCCAGTTCCTTCTTTCTGGGGATTGGTGAA
CTCAGCTTGGAATCTTTGCTCTGTGGGGAACCGGCAGTCGCCAGTCAACATAGAGACCAGTCA
CATGATCTTCGACCCCTTTCTGACACCTCTTCGCATCAACACGGGGGGCAGGAAGGTCAGTGG
GACCATGTACAACACTGGAAGACACGTATCCCTTCGCCTGGACAAGGAGCACTTGGTCAACAT
ATCTGGAGGGCCCATGACATACAGCCACCGGCTGGAGGAGATCCGACTACACTTTGGGAGTGA
GGACAGCCAAGGGTCGGAGCACCTCCTCAATGGACAGGCCTTCTCTGGGGAGGTGCAGCTCAT
CCACTATAACCATGAGCTATATACGAATGTCACAGAAGCTGCAAAGAGTCCAAATGGATTGGT
GGTAGTTTCTATATTTATAAAAGTTTCTGATTCATCAAACCCATTTCTTAATCGAATGCTCAA
CAGAGATACTATCACAAGAAATAACATATAAAAATGATGCATATTTACTACAGGGCTTAATAT
AGAGGAACATATCCAGAGACCTCTAGTTTTCATCACTTACGATGGGTGCGATGACTATCCCACC
CTGCTATGAGACAGCAAGTTGGATCATAATGAACAAACCTGTCTATATAACCAGGATGCAGAT
GCATTCTTGGCCTGCTCAGCCAGAACCAGCCATCTCAGATCTTTCTGAGCATGAGTGACAA
CTTCAGGCCTGTCCAGCCACTCAACAACCGCTGCATCCGCACCAATATCAACTTCAGTTTACA
GGGGAAGGACTGTCCAAACAACCGAGCCCAAGAAGCTTCAGTATAGAGTAAATGAATGGCTCCT
CAAG**TAG**GGAACAAAGCCAAGAAGAATCCCACCTCAGTGAAATGCTACAACCTGTGAATTGACG
TAACCTAGAATGTCCCCCTTCTTGCTTCTCTCTCTTCTTTCCCCCAAGCCTCATTCAATTCTT
GGGATTGGCCCTTTCTTCATGAAAAGTGTCTGCGAAACCATGGCAGAGGAATACATCTCTCAC
ACATACTCACAAACACACACACAAGCACTTGCACATACATACAAACACATGCAAACATACCTA
CACACACACACTCTCTTACAACCTCCATCATGGGAAGTCAAGTTTCAGAAACAAAAGTCTCAT
TCATAAGAGGTCTTAGAAGAAAATAACCAGTTAACCTGATTTCAATTTTGATACCGTTTTCTT
GAACTAATAAATCTACCCAATGAGACTTTTTCAGCCTTTGTACATACAAAATTCTTCCAAAAGA
GAGAGGAGAAAATACAGCTCTGATGGCATCAAACGGACTTTGCATCAAGTAATTCAGATAGT
GTCCTAGGATCCTTTGAGGGTGCTGGTAGCAGGTGAGCAGGACAAAGTTGACCAAGGACACTT
ATTTCTAGATTATGATTCTTCTGTTTACTCAACAATTTACAAAGAAAAAAGGACAGACATTG
AAGAGCTACACATTGTATATATATCACCACAGACTATAAGGAAATGGAATTATTTCCCTCTTT
GTCACATATCTGTAGTAGGATTTGCCAAGATCAGAAATGATCCATTTGCTGTTTCTTGTTTTT
CAAAGGTCATACATTGTGTTTGGTTATTGTTACCAGCTCAATAAATGTGTTTAAACGAGTTAAT
TTCATTTTTCTGGCTTTGGTCTGTTCTCCTTCCTTACAGGCTAAGCCCTGGCTCCATGCAACT
GCATTCTTTGATTTCACTTGTTTCCTTCATCTACATGTTTTGTTCAATTTGCAGCCAGTTTTTAC
TGAGTTTGTGGCAATCAGGAATGCATTTGCTAAGCAAGTATGACTTTAATTCCACTCCATGGC
TCAATCATTACATGAGGTGAGCTTCAGCCTGAGATAGCAGGCGACAGACTTCTTGCGTTTCA
AACTGCCATGCCCCCTGTGATGCTCCCGTGAAGGAATGCACTTTGCCTTGTAAGTTCCTGG
GAAAGGGGTATGTTTTCTCTCCAGGTGCAGCCAGATCTCACAAAGTACAAAACGAATGCCTTT
CTTTCTTGTTTATAATGGTCACTCACTGTGTTTGGTTACTGTCAAGAAATCAATAAATGTGT
TTAACAAGTTA

34/550

FIGURE 34

MEIVWEVLFLLLQANFIVCISAQQNSPKIHEGWWAYKEVVQGSFVPVPSFWGLVNSAWNLC SVG
KRQSPVNIETSHMIFDPFLTPLRINTGGRKVS GMTYNTGRHVSLRLDKEHLVNISGGPMTYSH
RLEEIRLHFGSEDSQGSEHLLNGQAFSGEVQLIHYNHELYTNVTEAAKSPNGLVVVSIFIKVS
DSSNPFLNRMLNRDTITRITYKNDAYLLQGLNIEELYPETSSFITYDGSMTIPPCYETASWII
MNKPVYITRMQMHSRLRLLSQNQPSQIFLSMSDNFRPVQPLNNRCIRTNINFSLQKDCPNNRA
QKLQYRVNEWLLK

Important features:**Signal peptide:**

amino acids 1-20

Eukaryotic-type carbonic anhydrases proteins.

amino acids 126-162, 220-269, 43-91

N-glycosylation sites.

amino acids 116-119, 168-171, 302-305

35/550

FIGURE 35

GTCTGGAACCCCTCAGGCCACCCTCGGGAGTCCTGGGGTCCAGAGGGGTGTCCCTGTACCCCTTGACAC
ACAGGACCCTCACTCTGCAGGGATAAGCCAGCTGCGCCTGCAGCCTAGGGTGCCAAAGGAGGCTGCTGA
TTGTGGCCACAGCCTCATCTGAACGCCAGGAGACCAGGATACCGAGGCACCGGATCCCCCTCTCTGTG
CCCTGGGGAGCCCCAGTGCTGCCAGTACCCCCAGGGCTGAGGTCTGCGTCCCTAGTGGTGCAAGGCC
TGGTAGGACCACGGGGCAGGGAATGTGAGCGCCATCCGAGCTCACGGTGTCTGAGTCGCGGCTTCGT
GACTTTGGCAGGGGCTCCGGACCAGTGACCCCCAGTCAAACCCAGAGGGTCTTGGGCGGCAGCGACGA
AGGAGGTATTAGGCTCCAGGCCAGGTGGGGCCGGACGCCCCAGCCATCCACCATGGTGGTGGCACA
CCCCACCGCCACTGCCACCACCACGCCCCACTGCCACTGTACGGCCACCGTTGTGATGACCACGGCCA
CCATGGACCTGCGGGACTGGCTGTTCTCTGCTACGGGCTCATCGCCTTCCTGACGGAGGTCATCGAC
AGCACCACCTGCCCCCTCGGTGTGCCGCTGCGACAACGGCTTCATCTACTGCAACGACCGGGGACTCAC
ATCCATCCCCGCAGATATCCCTGATGACGCCACCACCCTCTACCTGCAGAACCAACCAGATCAACAACG
CCGGCATCCCCCAGGACCTCAAGACCAAGGTCAACGTGCAGGTTCATCTACCTATACGAGAATGACCTG
GATGAGTTCCCCATCAACCTGCCCCGCTCCCTCCGGGAGCTGCACCTGCAGGACAACAATGTGCGCAC
CATTGGCAGGGACTCGCTGGCCCGCATCCCGCTGCTGGAGAAGCTGCACCTGGATGACAACCTCCGTGT
CCACCGTCAGCATTGAGGAGGACGCTTCGCGCAGACAACAGCTCAAGCTGCTCTTCCTGAGCCGG
AACCACCTGAGCAGCATCCCCCTCGGGGCTGCCGCACACGCTGGAGGAGCTGCGGCTGGATGACAACCG
CATCTCCACCATCCCGCTGCATGCCTTCAAGGGCCTCAACAGCCTGCGGCGCTGGTGCTGGACGGTA
ACCTGCTGGCCAACCAGCGCATCGCCGACGACACCTTCAGCCGCTACAGAACCTCACAGAGCTCTCG
CTGGTGCGCAATTGCTGGCCGCGCCACCCCTCAACCTGCCCAGCGCCACCTGCAGAAGCTCTACCT
GCAGGACAATGCCATCAGCCACATCCCTACAACACGCTGGCCAAGATGCGTGAGCTGGAGCGGCTGG
ACCTGTCCAACAACAACCTGACCACGCTGCCCCGCGGCTGTTTCGACGACCTGGGGAACTGGCCAG
CTGCTGCTCAGGAACAACCTTGGTTTTGTGGCTGCAACCTCATGTGGCTGCGGGACTGGGTGAAGGC
ACGGGCGGCGTGGTCAACGTGCGGGGCTCATGTGCCAGGGCCCTGAGAAGGTCCGGGGCATGGCCA
TCAAGGACATTACCAGCGAGATGGACGAGTGTTTTGAGACGGGGCCGAGGGCGGCGTGCCCAATGCG
GCTGCCAAGACCACGGCCAGCAACCACGCTCTGCCACCACGCCCCAGGGTTCCCTGTTTACCTCAA
GGCCAAAAGGCCAGGGCTGCGCCTCCCCGACTCCAACATTGACTACCCCATGGCCACGGGTGATGGCG
CCAAGACCCTGGCCATCCACGTGAAGGCCCTGACGGCAGACTCCATCCGCATCAGTGGAAGGCCACG
CTCCCCGCTCCTCTTTCCGGCTCAGTTGGCTGCGCCTGGGCCACAGCCAGCCGTGGGCTCCATCAC
GGAGACCTTGGTGCAGGGGGACAAGACAGAGTACCTGCTGACAGCCCTGGAGCCCAAGTCCACCTACA
TCATCTGCATGGTCACCATGGAGACCAGCAATGCCTATGTAGCTGATGAGACACCCGTGTGTGCCAAG
GCAGAGACAGCCGACAGCTATGGCCCTACCACCACACTCAACCAGGAGCAGAACGCTGGCCCCATGGC
GAGCTTGCCCTGGCGGGCATCATCGCGGGGAGTGGCTCTGGTCTTCTCTTCTGGTCTTGGGGG
CCATCTGCTGGTACGTGCACAGGCTGGCGAGCTGCTGACCCGGGAGAGGGCCTACAACCGGGGACG
AGGAAAAAGGATGACTATATGGAGTCAGGGACCAAGAAGGATAACTCCATCCTGGAAATCCGCGGCCC
TGGGCTGCAGATGCTGCCCATCAACCCGTACCGCGCCAAAGAGGAGTACGTGGTCCACACTATCTTCC
CCTCCAACGGCAGCAGCCTCTGCAAGGCCACACACCATTGGCTACGGCACACGCGGGCTACCGG
GACGGCGGCATCCCCGACATAGACTACTCTACACATTGATGCCCCGCCACCCGGGCTGCCCCGCTCA
GCCCCAGCTGCCCTGGCGTGGCCATGTGGCTTTGCCAGCCTGCTGCAATCCAAGAGAGCAAGGAAGA
GAAATTCCATGGGTGACTTTCCTCCGCAGAAAGCAAAGTTTGGGGAGGGCTGACGATTTTGTAGAACA
CAACAGTGACAATTTTTTTTAAAGAATAGAAGGCAGGAGGGGAATTGACATTGTTGAAGACATAA
TTTATACCAAGTTATGCCAGTTGGGGAGGGAAGGACTAAAATAATATTGCAGGCAGGGCTGGGTGG
GTTTTTTTTTTTCCCCCTGAACCTGGAAGGATACTACCTGTACAACATCTGTGGACACCTCATGCTCT
GTTCAAGGCCATCACAAAGGAACCGCCAGGGAGAAGCAGCCGGCTCTCAAAGCTCCCACGCAGCTCTC
CCGCCACTGGCCACTCGCTGGCGACCCGATGGAAGGTTTTTCAAGCTCCTCACAAAGGAGAGAGGGAAG
AAAAGATCTTTTGCCCTGGAGATATGGTCTGAAATCTCTCCCTGGCTTATTCCATACCATTTCCTT
TGCAGATTTGCAGAAACATGGCATCTTTCACTGCATTCTTTGAACAATCATGTAGTCGATTAAAAAAA
AAAACAACTTTTTTTTCCCTAGGCTGAAGCCCTCTTCAGTTCCATGCACCACGCTCCGTAGAAGCCCC
GGCGGAAGCCGTAGCTTTCCTGCCACCTGGAGGTGCATCTGTCTGCTGTCTATCCCTGTGCGGGTG
TCTCTAAGTACAGATGGGTAGATAGAGCCACATGCACGGTCTTACCCTTCTTCTTGGGTGAGTTCTT
ACCATTTCCTGAACAATAGAATTGTGAAAGTGTTAAAAA

36/550

FIGURE 36

MVVAHPTATATTTPTATVTATVMTTATMDLRDWLFLCYGLIAFLTEVIDSTTCPSVCRCNDG
FIYCNDRGLTSIPADIPDDATTLYLQNNQINNAGIPQDLKTKVNVQVIYLYENDLDEFPINLP
RSLRELHLQDNNVRTIARDSLARIPLLEKLHLDNSVSTVSIEEDAFADSKQLKLLFLSRNHL
SSIPSGLPHTLEELRLDDNRISTIPLHAFKGLNSLRRLVLDGNLLANQRIADDTFSRLQNLTE
LSLVRNSLAAPPLNLPSAHLQKLYLQDNAISHIPYNTLAKMRELERLDLSNNNLTTLPRLGFD
DLGNLAQLLLRNNPWFCGCNLMWLRDWVKARAAVNVNVRGLMCQGPEKVRGMAIKDITSEMDEC
FETGPQGGVANAAAKTTASNHASATTPOGSLFTLKAKRPGLRLPDSNIDYPMATGDGAKTLAI
HVKALTADSIRITWKATLPASSFRLSWLRLGHSPAVGSITETLVQGDKTEYLLTALEPKSTYI
ICMVTMETSNAAYVADETPVCAKAETADSYGPTTTLNQEONAGPMASLPLAGIIGGAVALVFLF
LVLGAICWYVHQAGELLTRERAYNRGSRKKDDYMESGTTKDNSILEIRGPGLQMLPINPYRAK
EEYVVHTIFPSNGSSLCKATHITIGYGTTRGYRDGGIPDIDYSYT

Important features of the protein:**Transmembrane domain:**

amino acids 552-573

N-glycosylation sites.

amino acids 249-252, 305-308, 642-645

Leucine zipper pattern.

amino acids 182-203, 299-320

Phospholipase A2 aspartic acid active site.

amino acids 57-67

FIGURE 37

[illegible]

38/550

FIGURE 38

MAEPGSHHLSARVRRRTERRIPRLWRLLLWAGTAFQVTQGTGPELHACKESEYHYEYTACDS
TGSRRWRVAVPHTPGLCTSLSDPVKGTECSFSCNAGEFLDMKDQSCKPCAEGRYSLGTGIRFDE
WDELPHGFASLSANMELDDSAESTGNCTSSKWVPRGDYIASNTDECTATLMYAVNLKQSGTV
NFEYYPDSSIIFFEFFVQNDQCQPNADDSRWMKTTEKGWEFHSVELNRGNVLYWRTTAFSVW
TKVPKPVLVNRNIAITGVAYTSECFPCPKGTYADKQSSFCCLKPANSYSNKGETSCHQCDDPK
YSEKSSSCNVRPACTDKDYFYHTACDANGETQLMYKWAKPKICSEDLEGAVKLPAAGVVKTH
CPPCNPFGFFKTNNSTCQPCPYGSYSNGSDCTRCPAGTEPAVGFEYKWWNTLPTNMETTVLSGI
NFEYKGMTGWEVAGDHIYTAAGASDNDFMILTLPVPGFRPPQSVMAADTENKEVARITFVFETL
CSVNCELYFMVGVNSRTNTPVETWKGSKGKQSYTYIIIEENTTSFTWAFQRTTFHEASRKYTN
DVAKIYSINVTNVMNGVASYCRPCALEASDVGSSTSCPAGYYIDRDSGTCHSCPPNTILKAH
QPYGVQACVPCGPGTKNNKIHSCLYNDCTFSRNTPTRTFNYNFSALANTVTLAGGPSFTSKGL
KYFHHFTLSLCGNQGRKMSVCTDNVTDLRIPEGESGFSKSIYAYVCQAVIIPPEVTGYKAGVS
SQPVSLADRLIGVTTDMTLDGITSAPAEFLHLESLGIPDVIFFYRSNDVTQSCSSGRSTTIRVR
CSPQKTVPGLSLLPGTCSGDTCDGCNHFHFLWESAAACPLCSVADYHAIVSSCVAGIQXTTYVX
REPKLCSGGISLPEQRTVICKTIDFWLKVGISAGTCTAILLTVLTCYFWKKNQKLEYKYSKLV
MNATLKDCDLPAADSCAIMEGEDVEDDLIFTSKKSLFGKIKSFTSKRTPDGFDSVPLKTSSGG
PMDL

Important features of the protein:**N-glycosylation sites:**

amino acids 153-156, 390-393, 391-394, 404-407, 544-547, 576-579,
672-675, 717-720, 947-950

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

amino acids 15-18, 563-566, 709-712

Casein kinase II phosphorylation sites:

amino acids 42-45, 59-62, 81-84, 146-149, 168-171, 282-285, 331-
334, 340-343, 431-434, 449-452, 465-468, 523-526, 557-560, 761-
764, 780-783, 835-838, 860-863, 893-896, 949-952

Tyrosine kinase phosphorylation sites:

amino acids 50-56, 109-116

N-myristoylation sites:

amino acids 77-82, 88-93, 152-157, 268-273, 288-293, 320-325,
400-405, 405-410, 414-419, 463-468, 599-604, 616-621, 634-639,
644-649, 839-844, 874-879, 912-917, 916-921

Amidation site:

amino acids 707-710

Cell attachment sequence:

amino acids 162-164

39/550

FIGURE 39

GGGAAGGGGTTCTGGGCTGCCGCAGGCACACAGGCCAGAGCTTCGTGGATACCTGCAGGGCCC
AAAGGTCCCTCCCTGTTTGAAGAGTGAGTGATGGCTATGAGGTAGCGGCCAGGCTGATCACC
CCTGCGTTGGCTGGAGGCAGAATTCTGTAAATCCTCGCCAAGTCTTTCTCCAGGCCACTGGTT
AGCTCATCTCAGCCTCCTCTGGGAGCATCAACACCAACATGGCACAGGGGACTGCAGTGGTGT
GCTTTGGACCTGTGTACCCACCCAAGGCTAAAGGCAGAGCCAGGTGACTTTGCGGGGGTCTCT
TCTCTAGGATTATCTGTACTTCCCCTCTGTCTCTTTTACTACGGGAGATCGAGCTAGCTATA
ACCCACCTTCTTTTCATGAGAACCACACTAAATTGCAAAAATTATCCCAGTGCTGGAGGAGGGC
AGCAGGTTGAGATTATGTTGGCAGGAAGAATGTTGGCATTGATTGGCACGCAGGGGACGAGAG
CTGCTTTGTGCTTTAAAGGAGCCAAGTTACACCCTGTTTAAACCCTGCCTTCAAAGGGACGACT
CTGTAAGATTCTCTGCTACTTATTCAAGTTGACACGATGCCCTTCACACTCCACCTGAGGTCC
CGCCTTCCCTCTGCCATAAGGAGTTTGATTCTACAAAAGAAACCAACATCAGAAATACATCC
AGCATGGCTGGAGAGCTCCGACCAGCCAGCCTGGTGGTCTGCCAGGTCCCTTGCTCCAGCT
TTTGAAGATTCTGCCAGGTCAACACTGGTCTCTACCCCTGCTGGGCCAGAGTGAGCCAGAA
AAGTGGATGCTGCCCCCTCAAGGTGCTATCTCAGAGACCAGGATGGGCCATCCCCAGTTCTGG
AAATACGAGTTCGGTGCCTGCACCGGTAGCCTGGCTTCGCTGGAGCAGTACTCGGAGCAGCTG
AAGGACATGGTGGCCTTCTTCCCTGGGCTGCAGCTTCTCCCTGGAGGAGGCCTTGGAGAAAGCG
GGGCTCCCCAGAAAGAGACCCAGCAGGTCACAGCCAGGCGGGTGCATACAAGACAACAGTGCCT
TGTGTTACCCATGCTGGCTTCTGCTGCCCTCTGGTGGTCACGATGAGGCCCATTCCTCAAGGAC
AAGCTGGAAGGGCTGGTGGGGCTGCTGCTCCCTCGGAGGTGAGCAGGGGCAACCTGTTTAC
ATGGGCGACCCAGAAGTGTGGGAATCAAAGAGCTTTCCAAACCTGCCTACGGGGATGCCATG
GTGTGTCCCCCAGGGGAGGTTCAGTGTTCTGGCCTTCTCCGCTGACCAGTCTCGGAGCTGTC
AGCAGCTGTGAGACCCCACTGGCTTTTGCCAGCATCCCAGGCTGCACAGTTATGACTGACCTG
AAGGATGCAAAGGCTCCACCTGGTTGTCTCACCCAGAGAGAATTCCAGAGGTCCATCACATT
TCCCAAGATCCTCTGCACTACAGCATCGCGTCAGTCTCTGCTTCTCAGAAGATCAGAGAACTA
GAGTCTATGATCGGCATAGACCCAGGGAACCGGGGGATTGGGCACCTGCTCTGTAAAGATGAG
CTGCTGAAGGCCTCTCTCTCGCTGTCCCATGCCGCTCAGTGCTCATCACCCTGGGTTCCTCC
ACACATTTCAATCATGAGCCTCCAGAAGAGACAGATGGCCCACCAGGAGCTGTTGCTCTGGTT
GCCTTCCTGCAGGCCTTGGAGAAGGAGGTGCGCATAATCGTTGACCAGAGAGCCTGGAACCTG
CACCAGAAGATTGTTGAAGATGCTGTTGAGCAAGGTGTTCTGAAGACGCAGATCCCGATATTA
ACTTACCAAGGTGGATCAGTGGAAGCTGCTCAGGCATTCTGTGCAAAAATGGGGACCCGAG
ACACCTAGATTTGACCACCTGGTGGCCATAGAGCGTGCCGGAAGAGCTGCTGATGGCAATTAC
TACAATGCAAGGAAGATGAACATCAAGCACTTGTTTGACCCATTGACGATCTTTTTCTTGCT
GCGAAGAAGATTCTGGAATCTCATCAACTGGAGTCGGTGATGGAGGCAACGAGCTTGGGATG
GGTAAAGTCAAGGAGGCTGTGAGGAGGCACATACGGCACGGGGATGTCATCGCCTGCGACGTG
GAGGCTGACTTTGCCGTGCTGCTGGTGTCTTAAGTGGGGAGGCTATGCCCTGGCCTGCGCA
CTCTACATCCTGTACTCATGTGCTGTCCACAGTCAGTACCTGAGGAAAGCAGTCGGACCCCTCC
AGGGCACCTGGAGATCAGGCCTGGACTCAGGCCCTCCCGTCGGTCATTAAAGGAAGAAAAATG
CTGGGCATCTTGGTGCAGCACAAAGTCCGGAGTGGCGTCTCGGGCATCGTGGGCATGGAGGTG
GATGGGCTGCCCTTCCACAACACCCACGCCGAGATGATCCAGAAGCTGGTGGACGTCACCACG
GCACAGGTGTAACCGTCCATGTTCCGTGTGAGCAGAGTCCCTACCAACGGGCAGGTCTGCATC
CGGGGAGAATGCAGCTGCTTCTGGCGACAATCCTGCTAGTAAACACTGGTCTTCGGTGAGCAA
CGAACACTCGCTGGCCTGGGAACTGCATGCCACTTTCTGGGAGGGGTTAGTGCAGGTGCC
GTGGACAAAGGACAACATTTCTCTGGGGCTTTTAACTTTTATTCCTAAGACTCTAAAGGCGT
TGATTTCAACCTCCTTCACTCTGGCTTCTTCAGGCAACCCACGTGGTCTCTATGAGAACTCT
TCTCGACAGTTACTTATGGGGACACTTGTGAACAATTAAGTGGCAGGGCAGAGCATGAGAACA
AATATCCAGGCCATGTAGGATAGGATACTCCAGACTCCAGTCATCCTCCCCATCCATGGT
TTCTGTTACTCATGGTTTTCAGTTACTCATAGCCAACTGCAGACCGAAAATACTAAATGAAAAA
TTTCAGAAATAAACAACCTCTTAAGTTTTTAAAAAAA

40/550

FIGURE 40

MPFTLHLRSRLPSAIRSLILQKKPNIRNTSSMAGELRPASLVVLPRSLAPAFERFCQVNTGPL
PLLGQSEPEKWMLPPQGAISETRMGHPQFWKYEFGACTGSLASLEQYSEQLKDMVAFFLGCSF
SLEEAL EKAGLPRRDPAGHSQAGAYKTTVPCVTHAGFCCPLVVTMRPIPKDKLEGLVRACCSL
GGEQGPVHMGDPELLGIKELSKPAYGDAMVCPGEPVFWPSPLTSLGAVSSCETPLAFASI
PGCTVMTDLKDAKAPPGCLTPERIPEVHHISQDPLHYSIASVSASQKIRELESMIGIDPGNRG
IGHLLCKDELLKASLSLSHARSVLITTGFPTHFNHEPPEETDGPPGAVALVAFLOALEKEVAI
IVDQRAWNLHQKIVEDAVEQGVLTQIPILTYQGGSV EAAQAF LCKNGDPQTFRFDHLVAIER
AGRAADGNYYNARKMNIKHLVDPIDDLFLAAKKIPGISSTGVGDGGNELGMGKVKEAVRRHIR
HGDVIACDVEADFAVIAGVSNWGGYALACALYILYSCAVHSQYLRKAVGPSRAPGDQAWTQAL
PSVIKEEKMLGILVQHKVRSGVSGIVGMEVDGLPFHNTHAEMIQLVDVTTAQV

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 358-378, 517-539

N-glycosylation site.

amino acids 28-32

Tyrosine kinase phosphorylation site.

amino acids 444-452

N-myristoylation site.

amino acids 98-104, 102-108, 123-129, 149-155, 181-187, 190-196,
238-244, 308-314, 399-405, 413-419, 448-454, 477-483, 482-488,
487-493

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 233-244, 531-542

41/550

FIGURE 41

CTTTCCTGTTTTATCCGCAGCCCTTTTCTTCTTTGAGTTAGTAAAGATTTATTCTGTAACCTG
ACACTCATCTGGCCCTTTGCAGTTTGCCAGCCATATTCCCATGTGATTTCCCACTGGATCCAG
GCCCCATCCGGCTGGCAGGAGGGGGCTCTGACGTACAGGTGGAAATCAGAAGTCTGTGAGA
GCGCGGGAGTG CATGGCAGCTCTGGGTCCCAGACCTGGCCCGACCCCTCTGCTTCACCTECAG
CTCTGCTGCTCCTCTACTCTTGGGTGAGATCCCTTTGGAGCCACAGCGAGGAACCCTGTGGT
CCTCAGGCAGGTGTACCTTGAGTCAGCCAGGAGCCCTCTTTTCCTGTGTCAAAGCCTGCCCTC
GGGCTCTGCTCACCTCTGGTGACCCTCCAAGATGCCCTGCCCTCAGTTTCCCCTCATGATCT
GGCCTCTGCCCCCTTCTCTAGCCACAGCCTCTAGTACACTTTAGCAATACCACCAGACTAGTT
AGAGTTCCCCACTCACCAAGCAAGACATGCAGTTTCATGCCTCTGTGCCTTCGCTCATGCTGT
TTCTTCCGACTGGAATGCCTTCCCCTGCTCCTCCTGCCTTGTCTGCCTGGCAAGTTCATCTCT
CACGATCCCCCTCAAAGGCCCCCTCCTCCAGGAAGGCAACCCCTGTGCCCTCCCCTCCAGGCT
ACCTCTGCACTTTGTCAATGCTTCTTGTGGCACTTATCACACTGTATTTTACTTGTTTACA
TGTTTGTCTCCCCTTCTAGACTGTGAATCCTTAAGGGCATGGACTGTATCTTATGCATCTCTG
TATTTCTGCGCCTAGCACGGTGCCTAGCACACAGTAGGCGCTCAATAAATGTTGAATGAATGA
ATGATTT

42/550

FIGURE 42

MQFHASVPSLMLFLPTGMPSAPPALSAWQVHLRSRPPPPGRQPLCPSPPGYLCTLSMLL
LWHLSHCILLVYMFVSPSRL

Important features of the protein:

Signal peptide:

amino acids 1-22

Microbodies C-terminal targeting signal.

amino acids 81-83

43/550

FIGURE 43

GTTTCCAACAAGGATGATATGAAGACTTCCCTGAAGAAAGTTGTGAAGGGACCTCCTACGAGA
TGATGATGCAGTGTGTGTCCCGCATGTTGGCCCAACCCCTGCATGTCATCTCAATGCGCTGCA
TGGTCCAGTTTGTGGGACGGGAGGCCAAGTACAGTGGTGTGCTGAGCTCCATTGGGAAGATTT
TCAAAGAGGAAGGGCTGCTGGGATTCTTCGTTGGATTAATCCCTCACCTCCTGGGCGATGTGG
TTTTCTTGTGGGGCTGTAACTGCTGGCCCACTTCATCAATGCCTACCTGGTGGATGACAGCT
TCAGCCAGGCCCTGGCCATCCGGAGCTATACCAAGTTCGTGATGGGGATTGCAGTGAGCATGC
TGACCTACCCCTTCTGCTAGTTGGCGACCTCATGGCTGTGAACAACCTGCGGGCTGCAAGCTG
GGCTCCCCCTTACTCCCCAGTGTTCAAATCCTGGATTCACTGCTGGAAGTACCTGAGTGTGC
AGGGCCAGCTCTTCCGAGGCTCCAGCCTGCTTTTCCGCCGGGTGTCATCAGGATCATGCTTTG
CCCTGGAGTAACCCTGAATCATCTAAAAACACGGTCTCAACCTGGCCACTGTGGGTGAGGCCT
GACCACCTTGGGACACCTGCAAGACGACTCCAACCCAACAACAACCAGATGTGCTCCAGCCCA
GCCGGGCTTCAGTTCCATATTTGCCATGTGTCTGTCCAGATGTGGGGTTGAGCGGGGGTGGGG
CTGCACCCAGTGGATTGGGTCACCCGGCAGACCTAGGGAAGGTGAGGCGAGGTGGGGAGTTGG
CAGAATCCCCATACCTCGCAGATTTGCTGAGTCTGTCTTGTGCAGAGGGCCAGAGAATGGCTT
ATGGGGGGCCAGGTTGGATGGGGAAAGGCTAATGGGGTCAGACCCACCCCGTCTACCCCTCC
AGTCAGCCCAGCGCCCATCCTGCAGCTCAGCTGGGAGCATCATTCTCCTGCTTTGTACATAGG
GTGTGGTCCCCCTGGCACGTGGCCACCATCATGTCTAGGCCTATGCTAGGAGGCAAATGGCCAG
GCTCTGCCTGTGTTTTTCTCAACACTACTTTTCTGATATGAGGGCAGCACCTGCCTCTGAATG
GGAAATCATGCAACTACTCAGAATGTGTCTCCTCATCTAATGCTCATCTGTTTAATGGTGAT
GCCTCGCGTACAGGATCTGGTTACCTGTGCAGTTGTGAATACCCAGAGGTTGGGCAGATCAGT
GTCTCTAGTCCTACCCAGTTTTAAAGTTCATGGTAAGATTTGACCTCATCTCCCGCAAATAAA
TGTATTGGTGATTGGAATAAAAAAAAAAAAAA

44/550

FIGURE 44

MMMQCVSRMLAHPLHVISMRCMVQFVGREAKYSGVLSSIGKIFKEEGLLGFFVGLIPHLLGDV
VFLWGCNLLAHFINAYLVDDSFSQLAIRSYTKFVMGIAVSMLTYPFLLVGDLMAVNNCGLQA
GLPPYSPVFKSWIHCWKYLSVQGQLFRGSSLLFRRVSSGSCFALE

Important features of the protein:

Signal peptide:

amino acids 1-18

Transmembrane domains:

amino acids 51-72, 97-114

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 160-163

N-myristoylation sites.

amino acids 34-39, 100-105, 123-128, 165-170

45/550

FIGURE 45

GCTCACTCTTTGGGTCCACACTGCCTTTATGAGCTGTAACACTCACTGGGAATGTCTGCAGCT
TCACTCCTGAAGCCAGCGAGACCACGAACCCACCAGGAGGAACAAACAACTCCAGACGCGCAG
CCTTAAGAGCTGTAACACTCACCGCGAAGGTCTGCAGCTTCACTCCTGAGCCAGCCAGACCAC
GAACCCACCAGAAGGAAGAACTCCAAACACATCCGAACATCAGAAGGAGCAAACCTCGTGACA
CGCCACCTTTAAGAACCGTGACACTCAACGCTAGGGTCCGCGGCTTCATTCTTGAAGTCAGTG
AGACCAAGAACCCACCAATTCCGGACACGGCAAAGTAACATCCTAGAC**ATGGCTTT**AGAGATC
CACATGTCAGACCCCATGTGCCTCATCGAGAAGTTTAATGAGCAGCTGAAGGTTAATCAGGAA
GCTTTGGAGATCCTGTCTGCCATTACGCAACCTGTAGTTGTGGTAGCGATTGTGGGCCCTCTAT
CGCACTGGCAAATCCTACCTGATGAACAAGCTGGCTGGGAAGAACAAGGGCTTCTCTGTTGCA
TCTACGGTGCACTCTCACACCAAGGGAATTTGGATATGGTGTGTGCCTCATCCCAACTGGCCA
AATCACACATTAGTTCTGCTTGACACCGAGGGCTGGGAGATGTAGAGAAGGCTGACAACAAG
AATGATATCCAGATCTTTGCACTGGCACTCTTACTGAGCAGCACCTTTGTGTACAATACTGTG
AACAAAATTGATCAGGGTGCTATCGACCTACTGCACAATGTGACAGAACTGACAGATCTGCTC
AAGGCAAGAACTCACCTGACCTTGACAGGGTTGAAGATCCTGCTGACTCTGCGAGCTTCTTC
CCAGACTTAGTGTGGACTCTGAGAGATTTCTGCTTAGGCCTGGAAATAGATGGGCAACTTGTG
ACACCAGATGAATACCTGGAGAATTCCTAAGGCCAAAGCAAGGTAGTGATCAAAGAGTTCAA
AATTTCAATTTGCCCCGTCTGTGTATACAGAAGTTCTTTCCAAAAAAGAAATGCTTTATCTTT
GACTTACCTGCTCACCAAAAAAGCTTGCCCAACTTGAAACACTGCCTGATGATGAGCTAGAG
CCTGAATTTGTGCAACAAGTGACAGAATTCTGTTCTTACATCTTTAGCCATTCTATGACCAAG
ACTCTTCCAGTGGCATCATGGTCAATGGATCTCGTCTAAAGAACCTGGTGCTGACCTATGTC
AATGCCATCAGCAGTGGGGATCTGCCTTGATAGAGAATGCAGTCCTGGCCTTGGCTCAGAGA
GAGAACTCAGCTGCAGTGCAAAAGGCCATTGCCCACTATGACCAGCAAATGGGCCAGAAAGTG
CAGCTGCCCATGGAAACCTCCAGGAGCTGCTGGACCTGCACAGGACCAGTGAGAGGGAGGCC
ATTGAAGTCTTCATGAAAACTCTTTCAAGGATGTAGACCAAAGTTTCCAGAAAGAATTGGAG
ACTCTACTAGATGCAAAACAGAAATGACATTTGTAAACGGAACCTGGAAGCATCCTCGGATTAT
TGCTCGGCTTTACTTAAGGATATTTTTGGTCTCTAGAAGAAGCAGTGAAGCAGGGAATTTAT
TCTAAGCCAGGAGGCCATAATCTCTTCATTAGAAAACAGAAGAACTGAAGGCCAAAGTACTAT
CGGGAGCCTCGGAAAGGAATACAGGCTGAAGAAGTTCTGCAGAAATATTTAAAGTCCAAGGAG
TCTGTGAGTCATGCAATATTACAGACTGACCAGGCTCTCACAGAGACGGAAAAAAGAAGAAA
GAGGCACAAGTGAAAGCAGAAGCTGAAAAGGCTGAAGCGCAAAGGTTGGCGGCGATTCAAAGG
CAGAACGAGCAAATGATGCAGGAGAGGGAGAGACTCCATCAGGAACAAGTGAGACAAATGGAG
ATAGCCAAACAAAATTGGCTGGCAGAGCAACAGAAAATGCAGGAACAACAGATGCAGGAACAG
GCTGCACAGCTCAGCACAACATTCAGCTCAAATAGAAAGCCTTCTCAGTGAGCTCCAGCAC
GCCCAGAGGGCTGTAAATAACGATGATCCATGTGTTTTACTCT**TAA**AGTGCTAAATATGGGAGT
TTCCTTTTTTTACTCTTTGTCACTGATGACACAACAGAAAAGAACTGTAGACCTTGGGACAA
TCAACATTTAAATAAACTTTATAATTATTAAA

46/550

FIGURE 46

MALEIHMSDPMCLIENFNEQLKVNQEALEILSAITQPVVVVAIVGLYRTGKSYLMNKLAGKNK
GFSVASTVQSHTKGIWIWCVPHPNWPNTLVLLDTEGLGDVEKADNKNDIQIFALALLLSSTF
VYNTVKNIDQGAIDLLHNVTELTDLLKARNSPDLDRVEDPADSASFFPDLVWTLRDFCLGLEI
DGQLVTPDEYLENSLRPKQGSQDQRVQNFNLPRLCIQKFFPKKKCFIFDLPAHQKKLAQLETLP
DDELEPEFVQQVTEFCSYIFSHSMTKTLPGGIMVNGSRLKNLVLTYVNAISSGDLPCIENAVL
ALAQRENSAAVQKAIAHYDQQMGQKVQLPMETLQELLDLHRTSERAIEVFMKNSFKDVDQSF
QKELETLLDAKQNDICKRNLEASSDYCSALLKDI FGPLEEAVKQGIYSKPGGHNLFIQKTEEL
KAKYYREPRKGIQAEVLQKYLKSKESVSHAILQTDQALTETEKKKKEAQVKAEAEKAEQRL
AAIQRQNEQMMQERERLHQEQVRQMEIAKQNWLAEEQQMQEQMQEQAAQLSTTFQAQNRSL
SELQHAQRAVNNDPCVLL

Important features of the protein:**Transmembrane domains:**

amino acids 31-49, 114-131

N-glycosylation sites.

amino acids 90-94, 144-148, 287-291, 563-567

N-myristoylation sites.

amino acids 45-51, 283-289

Prenyl group binding site.

amino acids 583-588

ATP/GTP-binding site motif A (P-loop).

amino acids 45-53

47/550

FIGURE 47

CACTCATTTCATTCCAAAGGGTCTCTCAAGGCAATGGTAATGTGCAAGGAGGTGATACCTAAAT
GAATGACCAAAAGAACATGCTTCTGCTTTTGTGTGTCTCCTACATTTTAGACATTTGTTTGTT
TCTCTTGGTAGCCTTTAAATTCCTTGAAGCCCAGGACCATGTCTCACTTACCTTTGTGTTTCC
ACTAACTAGTCTACCTCCTGGAATTGGCAGATACTCAGTGAAAGCCTGTGAAATAAGTGATGT
CTATTTCTAGCATATTATTCTGAGATTTAATGATAGATTTAGTGATTGAATGAGATTTCCATT
TTCAAATACAGCAAAAGCATAACTATTTTCATTCATTCATATTTCATTCAACTTCATTCTCAAA
ATTAGGTCCTGAGTTAACTAATAATTACCTTTGAAATGTGTGGGTATTTGAGGCAATCAGGT
GGTGACATTGAGCTCTCAGCCAGAGTTTGTTCCTGGAATTGATTCAGTTCCATTGCATTGATT
TTTGTTCCTCAGAAGCCAAGGTTTCCCATGAAAAATCATTCCCCTTGAATTGGGCTGTGATTC
TTGCTGCGTTTAAGTAAAGGAAGCCTCTTGGTCTAGTTCTGCAAACTTACACACTGAACTGG
GACAAGTTTTTGTTCAGAGTAATGGCTGGGAAAAGAGGAACCTTTCATTTTATTCAGAAGTCA
AAAACAAAGGCCCTCCAGCCACCTGGAGATGTTTTGTTGCAGACACCAGCCTGGCTCTGTCTT
TATGCCTAACAAATTGAGCATCCAGTCTTCTTTGTGCTGGGACCATTGCTCAGCTCTGCAAGGG
GAAAAGAGGGAGAAAGCCAGAGCTGCCAGGCTTCTTGCCTGAGGCGGGGGAGGGTTCCTGG
GAAGCAGGTGCTCTCTGGCTTCTTGGTACGTGAGGCTCTCGGAGCTGCCTCTCCTCTGACCCT
CAGGTCCTCACCGAGTTTGCTCCAGGAGTATATTGAAAACATACCCAGTGCTCTCTCAAGCAC
CCACTGCTTAGAGGGCCAGATTTCTTTTCTTCTTTCCCTTGCAAGCTGGAGACTGCATCG
GGCATCTGGTGTCTTAACTAAACAGGAAAAGTAAAGGTCCACAGTGCTCATTGTGTAGA
CTAGCTGCCCTCCGATGGGTGCTCTGATTATCAGTGGTTCCAGTGCAAGGCCTGTCACTAAAC
AGGCCTCACTTCCTCCTTGGGGGCTTTCCCATGGGAGGTGTGGCTTTTTACTCTACATGGAAA
TGACTCTCTGCAGCCACAGAACACAGTCATTTTCTGAATTATCCCAGTCTCTCATGCGCCCTG
GATTCCTCCAGATGCCTTATATCTTGTGCAAAGTTGTCTAAAATTTGGTTCCCAGCTTCCA
AGCCTTGCCTTTTGGCCTTCCTGGAAGTATTTTTGTTGATGAGTCGTCTGTCAATTATTCTCTA
AAATGATTTGCTTTTTGTTTCTTTCATTCCTATTTCCACCCACATATACACACATGCTTCTT
AACTTAGGGGATTACATGCCAATAAATCTATTGTTGAAAATGCACTAATACTATCGCAAAGAC
GAAAATTCACAGGCTGAACCGTTGTAAGTCCATATGCTCCTCAACTTACATGTGTGATGGAGT
TATGCCCAAATAAGTCCATCGTCAAGTTGAAAAATCAAATCAAGCCATCTTAGGTTGAGGAC
CATTTGTTTGTACCTCCAAAGATGTCATATCTTTAAACATACTCCCTAGCTTTTCTTTTACT
TTTTATTTTGAAGTAATTATAGAATCACAGAAAGTTGCAAAAAA

48/550

FIGURE 48

MGALIISGSSAGPVTQASLPPWGLSHGRCGFLLYMENTLCSHRTQSFSELSQSLMRPGFLQM
PYISCAKLSKIWFPAKPCLLAFLEVFLLSRSLFSKMICFLELSFLFPPHIYTHAS

Important features of the protein:

Signal peptide:

amino acids 1-41

Transmembrane domain:

amino acids 88-107

Casein kinase II phosphorylation site.

amino acids 47-50

N-myristoylation site.

amino acids 24-29

49/550

FIGURE 49

GGCTTCTACAGTCCACAACACCCACCAGCCCCAGGCCCAGCAGAATGAGCCCAGTGAGTGCCGGGGCTCCAGTT
TGGCTGTTGCTATGACAACGTGGCCACTGCAGCCGGTCTCTTGGGGAAGGCTGTGTGGGCCAGCCCAGCCATGC
CTACCCCGTGCGGTGCCTGCTGCCAGTGCCCATGGCTCTTGTGCAGACTGGGCTGCCCGCTGGTACTTCTGTGC
CTCTGTGGGCCAATGTAACCGCTTCTGGTATGGCGGTGCCATGGCAATGCCAATAACTTTGCCTCGGAGCAAGA
GTGCATGAGCAGCTGCCAGGGATCTCTCCATGGGCCCCGTCGTCCCCAGCTGGGGCTTCTGGAAGGAGCACCCA
CACGGATGGTGGCGGCAGCAGTCTGCAGGCGAGCAGGAACCCAGCCAGCAGGACAGGGGCCGCGGTGCAGAG
AAAGCCCTGGCCTTCTGGTGGTCTCTGGCGGCAAGACCAACAGCCTGGGCCAGGGAGGCCCCCACACCCAGGC
CTTTGGAGAATGGCCATGGGGGCAGGAGCTTGGGTCCAGGGCCCTGGACTGGGTGGAGATGCCGGATCACACG
GCCACCTTCCACAGCTCCTCTACAGATCTCACTTCCCACCTCTCCAGGATTAGCTTGGCAGGTGTGGAGCCCT
CGTTGGTGCAGGCAGCCCTGGGGCAGTTGGTGGCGCTCTCTGCTCAGACGACACTGCCCGGAATCCCAGGCTG
CCTGGCAGAAAGATGGCCAGCCCATCTCTTGACAGGCACAGGCTGCAGTTCCGACGGATCCCTGATCATCCACC
CCCTGCAGGCAGAGGACGCGGGCACCTACAGCTGTGGCAGCACCCGGCCAGGCCGCGACTCCAGAAGATCCAAC
TCCGATTATAGGGGGTGACATGGCCGTGCTGTCTGAGGCTGAGCTGAGCCGCTTCCCTCAGCCCAGGGACCCAG
CTCAGGACTTTGGCCAAGCGGGGGCTGCTGGGCCCCCTGGGGGCCATCCCCCTTTCACACCCACAGCCTGCAAACA
GGCTGCGTTTGGACCAGAACCAGCCCCGGGTGGTGGATGCCAGTCCAGGCCAGCGGATCCGGATGACCTGCCGTG
CCGAAGGCTTCCCGCCCCCAGCCATCGAGTGGCAGAGAGATGGGCAGCCTGTCTTCTCCCAGACACCAGCTGC
AGCCTGATGGCTCCCTGGTCAATAGCCGAGTGGCTGTAGAAGATGGCGGCTTCTACACCTGTGTCGCTTCAATG
GGCAGGACCGAGACAGCGATGGGTCCAGCTCAGAGTTCTGGGGGAGCTGACAATCTCAGGACTGCCCCCTACTG
TGACAGTGCCAGAGGGTGATACGGCCAGGCTATTGTGTGTGGTAGCAGGAGAAAGTGTGAACATCAGGTGGTCCA
GGAACGGGTACCTGTGCAGGCTGATGGCCACCGTGTCCACCAGTCCCCAGATGGCAGCTGCTCATTTACAAC
TGCGGGCCAGGGATGAGGGCTCCTACATGTGCAGTGCCACCAGGGGAGCCAGGCAGTCAGCCGACGACCCGAGG
TGAAGGTGGTCTCACCAGCACCCACGCCCCAGGCCAGGGACCCCTGGCAGGGACTGCGTCGACCAGCCAGAGCTGG
CCAATGTGATTTGATCCTGCAGGCCAGCTTTGTGGCAATGAGTATTACTCCAGCTTCTGCTGTGCCAGCTGTT
CACGTTTCCAGCCTCAGCTCAGCCCATCTGGCAGTAGGGATGAAGGCTAGTTCCAGCCCCAGTCCAAATAGTT
CATAGGGCTAGGGAGAAAGGAAGATGGAATCTTGGCTTCTCTCTGCTGGCTGGCAAAGGGAGTTATCTTCTGGAA
TACATTAGCTCTTTCAAAAACCCACCCAGTGTTTAGCCTCAACGGCAGCCAGTTACCAGCTTCTCTGTAGCCT
TCAGCAGTGTGTCATCTCTGACATAACACAGGCTGCTGTTTTCAAGAAGAGCAATCTGTTTGGATAAGAAAAA
CCTTTACTTTTACAGCTTCCCTTTTATAATTTGTTTACACAGGAATAGTTAAATGCATTTGTTTGTGTTTGGAG
ACGGAGTTTCACTCTTGTGCCCCAGGCTGGAGGGCAATGGCGGATCTCAGCTCACTGCAACCTCCGTCTCCTGG
GTTCTTGATTCTCTGTGTGAGCCTTCTGAGTAGCTGGGATTACAGATGCCTATCACCATGCCTGGGTAATTTTT
GTATTTTTTAGTTGAGATGGGGTTTCCGCATGTTGGCCAGGCTGGTCTCGAATCTTGACCTCAGATGATCTGCCC
GCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCCAGCCAGCCATCAATGCATTTTTTTTTATTTT
TTTTTTGAGACAGAGTTTCGCACTTCTTGCCAGGCTGGAGTACAATGGTGCGATCTTGGCTCACTGCAACCTCC
ACCTCCTGGGTTCAAGCGCTTCTCCAGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGTATGTGCCACCATGCCT
GGCTAATTTTTGTATTTTTTGGTGGAGACGGGGTTTCTCCATGTTGGTCAGACTGGTCTTGAATCCCCGACCTCAGG
TAATCCGCCCCGCTCCGCCTCCCAAATGCTGGGATTAGAGGTGTGAGCCACTGTGCCAGCCCATCAATGTGTT
TTAAAGCTAGCTGTGAGGTTCCACTTAATTTAAAGCTGGGCAGGGAGATGTGTAATGATTTCAAAGTTAACACC
TGTTTGTGTTTTTAAAGGGCATGCCAAGTCTGCTGTATCAGGGAAGTATTCTGTGCTAAAATCAGCGATGGTTCA
TTGCTCTAGTCTCTCTCACCCTTCTAGGCAGTGCATCAGTCAGCTCTAAATCTGGTGCAGAGGGTTAACAGCATA
ACCCTTGTGGCAAATGGAATAGATGTTAAGACCTCAAATAGGGATTTGGGATGAAACAGCTGCAGTTAGCACT
GTTATCTGAGCATGAAAGAACTGGAACGCTCCTTACCTCGAGATGTTGGACCTTGAAGCCCTCCTGAGGCCAAC
ATGCAAATCTGGCTGTGACGGTTCATCTGACACCTGTGTAAAGCTGACCAGCCTGCTCTGTACAGTGACAATGAG
GAGCCCCCTCTTCCCTTAAGTAGGAATCTGTGAAGCAAAATGTTTGCTGCCAAAGACAAATCAGACTGTGAGTCA
TTAAAAACAGCATTAGCAGGATGAGGATAGCAATGGGGAAGGGTTGTGGGCAATGCAGTAACAGGGAAATGGCTT
CAGAAATGGTTTGAAGTGAAGACAACATTCTTCACTCTCAGGACTTCTAATTCCTTGATGCTAAAAGAAGAGG
CATGGATTCTATGAGCTTCCAAGTCCCTTTCCACTTTAACCTTCTACAAATCTTTCAGAGGACTGCCTAGTAGCA
AAGGTTATTCTTGACACAGGAAAGACGGGCATTACAGGGACCAAGCTCTGAAAGGTGACTTTTATTACCAACA
CACTGGCTGGAAAAGGGACAAACCACATCACGGGTGAGTGATACTTCTCAGTCTTCTCTACTCATTCAACAAAGG
AAATGTGGGCTGGGGCAGAGGTCTTTTTTCAATTAATACTGGAAAAATATTGAAGAGCATCCATGTTCACTTATG
GCTGGTTTTGCTATAGAAATTGAAAATAAAGGCCACTTTTTTG

50/550

FIGURE 50

MGPVVPSLGLLEGAPTRMVAAAVLQASRNPASTGQGPRCRES PGLLVVSGGKTNSLGQGRPPT
PRPLENGHGGRSLGPGPLDWVEMPDHQRHPSTAPPTDLTSHLSRISLAGVEPSLVQAALGQLV
RLSCSDDTAPESQAAWQKDGQPISSDRHRLQFDGSLIIHPLQAEDAGTYSCGSTRPGRDSQKI
QLRIIGGDMAVLSEAELSRFPQPRDPAQDFGQAGAAGPLGAIPSSHQPANRLRLDQNQPRVV
DASPGQRIRMT CRAEGFPPPAIEWQRDGQPVSSPRHQLQPDGSLVISRVAVEDGGFYTCVAFN
GQDRDQRWVQLRVLGELTISGLPPTVTVP EGD TARLLCVVAGESVNIRWSRNLPLVQADGHRV
HQSPDGTLLIYNLRARDEGSYMC SAYQGSQAVSRSTE VKVVSPAPTAQPRDPGRDCVDQPELA
NCDLILQAQLCGNEYYS SFCCASCSRFQPHAQPIWQ

Important features of the protein:**Signal peptide:**

amino acids 1-16

Tyrosine kinase phosphorylation site.

amino acids 392-400

N-myristoylation sites.amino acids 9-15, 50-56, 112-118, 146-152, 173-179, 195-201,
220-226, 229-235, 280-286, 306-312, 336-342, 397-403**Myelin P0 protein.**

amino acids 153-182

51/550

FIGURE 51

CAGGCAGAAGCGAACAAAGACCCAGCAAGAGAAGGCAGAGGCTAAGACCCATCCCGTATCTGC
TCTCCTGAAATAATTCTGGAGTCATGCCTGAAATGCCAGAGGACATGGAGCAGGAGGAAGTTA
ACATCCCTAATAGGAGGGTTCTGGTTACTGGTGCCACTGGGCTTCTTGGCAGAGCTGTACACA
AAGAATTTTCAGCAGAATAATTGGCATGCAGTTGGCTGTGGTTTTCAGAAGAGCAAGACCAAAAT
TTGAACAGGTTAATCTGTTGGATTCTAATGCAGTTCATCACATCATTTCATGATTTTTCAGCCCC
ATGTTATAGTACATTGTGCAGCAGAGAGAAGACCAGATGTTGTAGAAAATCAGCCAGATGCTG
CCTCTCAACTTAATGTGGATGCTTCTGGGAATTTAGCAAAGGAAGCAGCTGCTGTTGGAGCAT
TTCTCATCTACATTAGCTCAGATTATGTATTTGATGGAACAAATCCACCTTACAGAGAGGAAG
ACATACCAGCTCCCCTAAATTTGTATGGCAAAACAAAATTAGATGGAGAAAAGGCTGTCCTGG
AGAACAATCTAGGAGCTGCTGTTTTGAGGATTCCTATTCTGTATGGGGAAGTTGAAAAGCTCG
AAGAAAGTGCTGTGACTGTTATGTTTGATAAAGTGCAGTTCAGCAACAAGTCAGCAAACATGG
ATCACTGGCAGCAGAGGTTCCCCACACATGTCAAAGATGTGGCCACTGTGTGCCGGCAGCTAG
CAGAGAAGAGAATGCTGGATCCATCAATTAAGGGAACCTTTCCTGCTGGCAATGAACAGA
TGACTAAGTATGAAATGGCATGTGCAATTGCAGATGCCTTCAACCTCCCCAGCAGTCACTTAA
GACCTATTACTGACAGCCCTGTCCTAGGAGCACAACGTCCGAGAAATGCTCAGCTTGACTGCT
CCAAATTGGAGACCTTGGGCATTGGCCAACGAACACCATTTTGAATTGGAATCAAAGAATCAC
TTTGGCCTTTTCCTCATTGACAAGAGATGGAGACAAACGGTCTTTCATTAGTTTATTTGTGTTG
GGTTCCTTTTTTTTTTTTAAATGAAAAGTATAGTATGTGGCACTTTTTTAAAGAACAAAGGAAATA
GTTTTGTATGAGTACTTTAATTGTGACTCTTAGGATCTTTCAGGTAAATGATGCTCTTGCACT
AGTGAAATTGTCTAAAGAACTAAAGGGCAGTCATGCCCTGTTTGCAGTAATTTTTCTTTTTTA
TCATTTTGTTTGTCCTGGCTAAACTTGGAGTTTGTAGTATAGTAAATTATGATCCTTAAATATT
TGAGAGTCAGGATGAAGCAGATCTGCTGTAGACTTTTCAGATGAAATTGTTTCATTCTCGTAAC
CTCCATATTTTCAGGATTTTTGAAGCTGTTGACCTTTTCATGTTGATTATTTTAAATTGTGTG
AAATAGTATAAAAATCATTGGTGTTTATTATTTGCTTTGCCTGAGCTCAGATCAAATGTTTG
AAGAAAGGAACCTTATTTTTTGCAAGTTACGTACAGTTTTTATGCTTGAGATATTTCAACATGT
TATGTATATTGGAACCTCTACAGCTTGATGCCTCCTGCTTTTATAGCAGTTTATGGGGAGCAC
TTGAAAGAGCGTGTGTACATGTATTTTTTTTCTAGGCAAACATTGAATGCAAACGTGTATTTT
TTAATATAAATATATACTGTCCTTTTCATCCCATGTTGCCGCTAAGTGATATTTTCATATGT
GTGGTTATACTCATAATAATGGGCCTTGTAAGTCTTTTCACCATTTCATGAATAATAATAATA
TGTACTGCTGGCATGTAATGCTTAGTTTTCTTGATTTTACTTCTTTTTTTTAAATGTAAGGACC
AACTTCTAACTAATTGTTCTTTTGTGCTTTAATTTTTTAAAAATTACATTCTTCTGATGTA
ACATGTGATACATACAAAAGAATATAGTTTAATATGTATTGAAATAAAACACAATAAAATT

52/550

FIGURE 52

MPEMPEDMEQEEVNIPNRRVLVTGATGLLGRAVHKEFQQNNWHAVGCGFRRARPKFEQVNLLD
SNAVHHIIHDFQPHVIVHCAAERRPDVVENQPDAAASQLNVDSAGNLAKEAAAVGAFLIYISSD
YVFDGTNPPYREEDIPAPLNLYGKTKLDGEKAVLENNLGA AVLRIPILYGEVEKLEESAVTVM
FDKVQFSNKSANMDHWQQRFPTHVKDVATVCRQLAEKRMLDPSIKGTFHWSGNEQMTKYEMAC
AIADAFNLPSSHLRPITDSPVLGAQRPRNAQLDCSKLETLGIGQRTPPFRIGIKESLWPFLIDK
RWRQTVFH

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 105-127

N-glycosylation site.

amino acids 197-201

N-myristoylation site.

amino acids 303-309

Short-chain dehydrogenases/reductases family proteins.

amino acids 18-30

53/550

FIGURE 53

TGGGCTCCCTCCAGCACTGCTGTTGCCTGCTGCCTAAGATGGGTGACACTTGGGCCAGCTTCCCTGGCCTGGGG
CAGCCACCCAGCAATGCTGCTGATCTCCCTCCTCTTGGCAGCCGGGTTGATGCACTCGGATGCCGGCACCAGCT
GCCCCGTCTTTGCACATGCCGTAACCAAGTGGTGGATTGTAGCAGCCAGCGGCTATTCTCGTGCCCCAGACC
TGCCAATGGACACCCGAAACCTCAGCCTGGCCCAACCGCATCACAGCAGTGCCGCTGGCTACCTCACATGCT
ACATGGAGCTCCAGGTGCTGGATTTCACAACAACCTCCTTAATGGAGCTGCCCCGGGGCCTCTTCTCCATGCCA
AGCGCTTGGCACACTTGGACCTGAGCTACACAATTTAGCCATGTGCCAGCCGACATGTTCCAGGAGGCCCATG
GGCTAGTCCACATCGACCTGAGCCACAACCCCTGGCTGCGGAGGGTGCATCCCCAGGCCTTTAGGGCCTCATGC
AGCTCCGAGACCTGGACCTCAGTTATGGGGCCTGGCCTTCTCAGCCTGGAGGCTCTTGAGGGCCTACCGGGGC
TGGTGACCTGCAGATCGGTGGCAATCCCTGGGTGTGTGGCTGCACCATGGAACCCCTGCTGAAGTGGAGGCGCCCCG
ACCGGATCCAGCGCTGTACAGCAGATTCTCAGCTGGCTGAGTGGCGGGGCCCTCTGAAGTCGAGGGCGCCCCG
TCTTCTCACTCACTGAGGAGAGCTTCAAGGCCTGCCACCTGACCTGACCTGGATGATTACCTATTTCATTGCGT
TCGTGGGCTTCTGGTCTCCATTGCTTCTGTGGCCACCACTTCTCTGGGCATCACTGCCAATGCTGCCACC
GCTGGAGCAAGGCCAGTGAAGAGGAAGAGATCTGACATGCTGCTCCTCTCATCCCTCCATGCTGCTGACCGCCACA
GCTGCTGGCCACCAGACGCCCTCCCTGATTGCTCACTCTGCTTCCATGGTGACCTGGCTGCCCTCAGTCATGGTTC
AAGCAAGGTGGGGACACTCATTTTGTATGAGCATCTGCTTGGGCCAGGCGGCACGCTAGGAATTGGGAACATCA
GATGAACTGACTCAGTCCCTGCCCTCAAGGCACTTCCCTCTGGTCAAGGAGAGAGATCCAAAACCTATTCCCTTT
AAGACTATATGTCAGGACTCTGAGCAGTCATTATGGAGGGCCAGAGGAGGAGCCATCATCTGTATCTAGCAATG
TCCATGAGAATTATAAGATTAGAGTGATTGTGAACCTGGGTGCATCAGGAAATATCTACTTTGTCAGGTAGGCAAA
GAAGGTGTCTGCACATGGCAGAGGCCAGAATATGCATAGTGTGCTGTGTTGAGAAGAGTGAACAGTTCTCTGGTC
ACTTACTTGTATAGAGGGGTGTGGCACAGAACTCAAACCTACCCCTCACCTCCTGACACCAAACTGTGAGCTC
TCAGCAATGCCAGCCACTGCCTACAGGAGTAAGAACCTCTATGACAGCCCTGGCCTCCTTCCACCAGCAGC
TACCAGGTGAGACCACCTCCAGTGACTGCCCCATATGACCAAAATGTCAACAGTTGGTGAGGTCCCAGGCAGCA
GGCTGAGGATGGACACTTTCAATGCCCTTGTCTGCTCCTCACTCAAGTTTGTCTCAGAAGAGAGAGGCAGGA
GGCCAGCAACTGGGGCAGCAAGAGTCTGGCACCTTGGGATCCTAATCATGTGACTGTCTTGGCACAGTGTCT
ATGCCACAGGGTCTCACCAGGAAAGTGCACTGTGGGCCACAGACCCACAGCCTGGCAGCACCAGAGCTAAAAGG
GGCAAAAGGCAGCAGCTTATGACCATATGAGGCTTTGCATTTTCTTCTAAGCAACTTACCCACGTTAAGCATGA
GGGTGAGAGAGCTATTAATACTAAGCCCTTGGCAGTGTCAAGTACTTTGAAAAGCTCTCTGCACAAACCATTTCC
CTTTGACACACACACACAAATCTTTGAGGTGAACCTGTTGTTGTTCCATTTTACGGATGAGGCAACTAAGGCT
CAGAGAGGTTAAAGTCACATGCCACTATGAGCAAGATAAAGTCTGTGCTCTTTCTACTGCCCCCAAGTTGGG
GAACATCACCATTCCCTCTAGAGTTATATAAATTCAAATCAACTAGAGCTGACAAAGTTCTCATAAGGTCCAG
GCACTCCTCTGGGCACTTTTATATCTATTGACTCACTTCTTTCAATTCTCACAGCAACCTGCCTGGTGGTTTTT
ATTATCCCCATTGACAGATGAATTAATCGTAGAGAGTTGAGTGACTTACCAAGGTTGTCTGGATAAGCCCTAG
AAGGAAGGCGGTAGGCAGCTCCATTCAAGGAACTGCATCTAATCAGTCAGTCAAAAATCAAGTAACCTTACGAG
CAAAGCACAATTATCATCATCGTGGTCTTCTTCATCAGTTTCTGTCAGCAGCATCATTTCTTCCCTCTATTGTT
CAGCACCGGATAGTTCATGAGTATTTTGCATCATTCTCCTTGACTTTTACATCCCTGTGCAGGAGGTAAATCA
AACATCAGTAATCCTGTTTTACAGATGGGGAAAAAGTCTCAAGGTTGGATATGACTTGCTATGTGGCAAGGTTG
GGGCTCAACCTTAACAGATTCTCTTCCAGTGCTTCTCAAGTGCTTGGGGAAGAGAATGCCTCAGAAGGCTGG
GTAGTGGGGCCTTGAATTCAGCATCCATGAATGCTAGTGGAATAAGCTAAATAGAAGGCAGCCAAACCATCT
GCTGTACAGATTGAACTATGCTCAGGTAGGGCAAATTGCAGGCTCTGAAACAGAGACTACACAGGTAAACCTG
AATAGGAGACTCCTGCTTTACAATGTGTAGATAAAACATCAGCAATGGTGGCCATGGTGGCAGTCATGTGAAAAG
TAAGATCTTTGGGAATCAAGAAAGGAAGCTGTGTTAACCACCTCCTGCTCAAGCCCTGCTGCGTGTGTGCAAGAG
ATACTAAGAGAGCAAGAAAGCTATAGGTGAGAACCTCTGCAGTTTAGGAGAAGAACATCAAGGCACAGTCCAACA
TGCTGATAAGTCTGGCCAGGAGGAGAAATTAACACAGGGGCTTTCCACACCTCCCTTGCCCCAAGCTCCAGCGGTA
TTCTATCAGCCCATCCTCCTGGAAAGCCTGAAAGGAATGAAGGAGGCTAATAAGTCATCTTCCAGGAAGGCATCC
CTCACTCGTGCTTCCCTGAGCTAGTCAACCAAGAGTCTTACAGAACTTTGCTAGACCTGAAGTACTTGAACCT
GTGTCCCTGAATCTTTCTTACAACATCTGGGACAAATCCCTGGTCTGTGACATCCGAAGCAGAACTGTGCCCT
GCTCTCTCCTTCTGTGATGACCAAGGATGGTGAACCTCAAGTTGTTCTCTACAAGCCAGGCCAGCAACCTAAATAC
TTGGAGAGGAACTTTAGAACTATAATCCTGACAAATAGAAAAGTTTCCATAGGGGCATACCAATACTAT
AATAACCTCCAGGAACCTATTGTTTGCAAAATGTAGTTAATATATTTTAAGATATATGCTTTTTTGCATAGGAC
TAGAACCAGAAAAGACACCAATGCCCCCTTGACATCAATGTCCTTTCTAGTGGGACAATTTGGTCTCCATTAAT
GCCAAACCTTTCTGAACAGGATACATGGCTTTTAAAGGACAGATGTTTCTCCTGCTGCTAGAAGTTCTCAGTTT
ACTAGAGCACAATGAGGAAAGTATTCAACCTCCCTACTGCCAAGGAATCCCTGCTTCTCCCCACCGCCATCAT
CTTGTCGAAGCTTATCAGAAGCAACCTTCTAGAGATAATCTAACAATCCTGATTAGAATTGCTCCCATATCCCTGG
TGACCACAGGCTTCATTCAAATTTGTCCAACTGGTTAACATGTATGTGATGGGGTATCTCTGCATCTGTATGTCT
GTCTGCGAGGTTCTTGTATATTGGCTGTCCGCTGACTTGGGACAGATCTCTCTAGAACTTGGGTTCAAGTTCTCT
GACATAGTCCACTCAGCCATAGGCTGAGTGGCTAAATATGCATAAATAAGCATGCCATAATAGGCATATATAGGT
TGGTGCAAAAGTAATTGCGGTTTTTGCATTAAAAATGATGGCAAAAATCCCAATTACTTTTGGCTCAATCTAAT
ATTACATTGCTTGATAGATTAAAGATGGAATCCACCAGGTTTAGGGTAGGACTGGATGCTCAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

54/550

FIGURE 54

MLLISLLLAAGLMHSDAGTSCPVLCTCRNQVVDCCSQRLFSVPPDLPM DTRNLSLAHN RITAV
PPGYLTCYME LQVLDLHNNSLMELPRGLFLHAKRLAHL DLSYNNFSHV PADMFQE AHGLVHID
LSHNPWLRRVHPQAFQGLMQLRDL DLSYGG LAFLSLEALEGLPGLVTLQIGGNPWVCGCTMEP
LLKWLNRNRIQRCTADSQ LAECRGPPEVEGAPLFSLTEESFKACHLTLT LDDYLFIAFVG FVVS
IASVATN FLLGITANCCHRWSKASEEEEEI

Important features of the protein:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 241-260

N-glycosylation sites.

amino acids 52-55, 81-84, 107-110

Tyrosine kinase phosphorylation site.

amino acids 148-154

N-myristoylation sites.

amino acids 11-15, 263-268

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 175-185

Leucine zipper pattern.

amino acids 77-98

55/550

FIGURE 55

GGCTGCGCCCAGGCCGGCGGGCCAGCAGCTGCGAACCGCCGGCGCACCACCTGTTTTCCGCGC
CCGGGGACTTCCCCGGCGGGGCTCAGAAGTGTGGGGTCGGTCGCTTGGCTTCCCCTGGCGTCA
GCGACCCAGGGTAACCTCCTCCACTGCTGCGTGCCGTGCAGGCCTGCCTGTGTGAGAGCCACG
TGTGCCGCGCTCTGGGCACAGCCTTGGAAAGTCAGGACCGCGACGGCAGCAGAGCAGAAACCT
TACAGAAACATGAAGCCCTCAACCATCTGCTACTCAGTTATTCGGGGCTGACGGCGGCTTCTA
GAACATCCAGGTGTTCTGCAGATGCGAGAACTCATCCTGTAGTCACCAGATGGAGTCCCAAAC
AGCCAAGCAGATGTAAGGCCTGTGCTGTGGCTCTGAGGCCCTGAATACAGAAGGGTCACTTTC
TTAGTGGCCAAAGAGCAGTTGTTGACATTGATGTCTAATTATTGAACACGACCAGTCATTTTA
CTGAGCTGCAGTGAGGAAACACTGACCATAGAAGATCAAGCCAAATGAGGGATTGCAAATTTT
CTGATTCTTTTGAATTAGGATTCAGATGGGGGCCTCATTTCTACAGCCCCCAACATTCCTAT
AGCCGTTATCACTGCCATCACCCTGCCACCAGCATCTTCTTGCAGATTCCACCCCTGCTCCC
CAGAGACTTCCTGCTTTGAAAGTGAGCAGAAAGGAAGCTCTCAGAAAAATCTCTAGTGGTGGC
TGCCGTGCTCCAGACAATCGGAATCCTGCCTCACCACCATGGGCTGGCTTTTTCTAAAGGT
TTTGTGGCGGGAGTGAGTTTCTCAGGATTTCTTTATCCTCTTGTGGATTTTTCATCAGTGG
GAAACAAGAGGACAGAAGCCAACTTTGTGATTATTTTGGCCGATGACATGGGGTGGGGTGA
CCTGGGAGCAAACCTGGGCAGAAACAAGGACACTGCCAACCTTGATAAGATGGCTTCGGAGGG
AATGAGGTTTGTGGATTTCATGCAGCTGCCTCCACCTGCTCACCTCCCGGGCTTCCTTGCT
CACCGGCCGGCTTGGCCTTCGCAATGGAGTCACACGCAACTTTGCAGTCACTTCTGTGGGAGG
CCTTCCGCTCAACGAGACCACCTTGGCAGAGGTGCTGCAGCAGGCGGGTTACGTCCTGGGAT
AATAGGCAAATGGCATCTTGGACACCACGGCTCTTATACCCCCAACTTCCGTGGTTTTGATTA
CTACTTTGGAATCCCATATAGCCATGATATGGGCTGTACTGATACTCCAGGCTACAACCACCC
TCCTTGTCAGCGTGTCCACAGGGTGATGGACCATCAAGGAACCTTCAAAGAGACTGTTACAC
TGACGTGGCCCTCCCTCTTTATGAAAACCTCAACATTGTGGAGCAGCCGGTGAACTTGAGCAG
CCTTGCCCAGAAGTATGCTGAGAAAGCAACCCAGTTCATCCAGCGTGCAAGCACCAGCGGGAG
GCCCTTCCTGCTCTATGTGGCTCTGGCCACATGCACGTGCCCTTACCTGTGACTCAGCTACC
AGCAGCGCCACGGGGCAGAAGCCTGTATGGTGCAGGGCTCTGGGAGATGGACAGTCTGGTGGG
CCAGATCAAGGACAAAGTTGACCACACAGTGAAGGAAACACATTCCTCTGGTTTACAGGAGA
CAATGGCCCGTGGGCTCAGAAGTGTGAGCTAGCGGGCAGTGTGGGTCCCTTCACTGGATTTTG
GCAAACCTCGTCAAGGGGGAAGTCCAGCCAAGCAGACGACCTGGGAAGGAGGGCACCGGGTCCC
AGCACTGGCTTACTGGCCTGGCAGAGTTCCAGTTAATGTACCAGCACTGCCTTGTTAAGCGT
GCTGGACATTTTTTCCAACCTGTGGTAGCCCTGGCCAGGCCAGCTTACCTCAAGGACGGCGCTT
TGATGGTGTGGACGTCTCCGAGGTGCTCTTTGGCCGGTCACAGCTGGGCACAGGGTGCTGTT
CCACCCCAACAGCGGGGCAGCTGGAGAGTTTGGAGCCCTGCAGACTGTCCGCCTGGAGCGTTA
CAAGGCCTTCTACATTACCGGTGGAGCCAGGGCGTGTGATGGGAGCATGGTGCCTGAGCTGCA
GCATAAGTTTCTCTGATTTTCAACCTGGAAGACGATACCGCAGAAGCTGTGCCCTAGAAAG
AGGTGGTGCAGGAGTACCAGGCTGTGCTGCCCCAGGTGAGAAAGTTCTTGCAGACGTCCTCCA
AGACATTGCCAACGACAACATCTCCAGCGCAGATTACACTCAGGACCCTTCAGTAACCTCCCTG
CTGTAATCCCTACCAAATTGCCTGCCGCTGTCAAGCCGCATTAACAGACCAATTTTATTCCAC
GAGGAGGAGTACCTGGAAATTAGGCAAGTTTGCTTCCAAATTTTCAATTTTACCCTCTTTACAA
ACACAGCCTTTAGTTTAGTCTTGGAGTTTAGTCTTGGAGTTAGCCTTGCATATCCCTTCTGTA
TCCTGTACCCCCCTCCACGCCGACCCGAGAGAGCTGAGCTGCGCTGGCTCTGGGCAGGGAGTG
TGCTTAAATGGGAAGCACACGGGCTTTGGAGTCAGGCACAGGTGCCAGCTCCAGCTTTTGAAC
TTGGGCAATTGTTTAACTAACCTGCAAGTTGATTTTGAAGGTTAAATAAAGGCATACATGAA
AATGCCTGGCAACTTTAAAAA

56/550

FIGURE 56

MGWLFLKVLLAGVSFSGFLYPLVDFCISGKTRGQKPNFVII LADDMGWGDLGANWAETKDTAN
LDKMASEGMRFVDFHAAASTCSPSRASLLTGRLGLRNGVTRNFAVTSVGGLPLNETTLAEVLQ
QAGYVTGIIGKWHLGHHGSYHPNFRGFDYYFGIPYSHDMGCTDTPGYNHPPCPACPQGDGPSR
NLQRDCYTDVALPLYENLNIVEQPVNLSLAQKYAEKATQFIQRASTSGRPFLLYVALAHMHV
PLPVTQLPAAPRGRSLYGAGLWEMDSLVGQIKDKVDHTVKENTFLWFTGDNGPWAQKCELAGS
VGPF TGFWQTRQGGSPAKQTTWEGGHRVPALAYWPGRVPVNVSTALLSVLDIFPTVVALAQA
SLPQGRRFDGVDVSEVLFGRSQPGHRVLFHPNSGAAGEFGALQTVRLERYKAFYITGGARACD
GSMVPELQHKFPLIFNLEDDTAEAVPLERGGAEYQAVLPEVRKVLADVLQDIANDNISSADYT
QDPSVTPCCNPYQIACRCQAA

Important features of the protein:**Signal peptide:**

amino acids 1-16

Transmembrane domain:

amino acids 353-373

N-glycosylation sites.

amino acids 117-120, 215-218, 356-359, 397-500

N-myristoylation sites.

amino acids 12-17, 33-38, 52-57, 97-102, 101-106, 113-118, 158-163, 328-333, 388-393, 418-423, 435-440, 436-441

Amidation site.

amino acids 382-385

Sulfatases signature 2.

amino acids 129-138

57/550

FIGURE 57

TGGACAAGACACCTCCAGGAGCCCAGCTCACAGCCACCGGTACCTTCTTCCAGGACAAGCTGG
GGGCCTCCATGGGCGCCTGAGGGCCAGGCGCCAGGGCCGTGGGCACGAGT**ATG**GTGAGACACC
AGCCCCTGCAGTACTACGAGCCACAGCTGTGCCTCTCCTGCCTCACGGGCATCTACGGCTGCC
GTTGGAAGCGCTACCAGCGCTCCCATGATGATACACACCGGGCACAGCGCCATTCCTGCATG
TGGGGGCTGTGGCAGCAGTCACCATGCTCTCCTGGATCGTGGCAGGACAGTTCGCCCCGTGCAG
AGCGGACCTCCTCCCAGGTGACCATTCTCTGTACCTTCTTCACCGTGGTGTTCGCCCTCTACC
TGGCCCCCTCTACCATCTCCTCTCCCTGCATCATGGAGAAGAAAGACCTCGGCCCCAAGCCTG
CTCTCATTGGCCACCGCGGGGGCCCCCATGCTGGCTCCAGAGCACACGCTCATGTCCTTCCGGA
AGGCCCTCGAGCAGAAGCTGTACGGGCTCCAGGCTGACATTACCATCAGCCTGGACGGCGTGC
CCTTCCTCATGCATGACACCACCTGCGGCGCACCACCAACGTGGAGGAGGAGTTCCCGGAGC
TGGCCCGCAGGCTGCCTCCATGCTTAAGTGGACACCCTGCAGAGACTCAACGCTGGCCAGT
GGTGAAGACTGACCCCTTCTGGACAGCGAGCTCCCTGTACCCTCCGACCACAGAGAGG
CCCAGAACCAGTCCATCTGCAGCCTGGCAGAGCTCCTGGAGCTGGCCAAGGGCAATGCCACAC
TGCTGCTCAACCTGCGTGACCGCCCCGGGAGCACCCCTACCGCAGCAGTTTTATCAACGTGA
CTCTGGAGGCCGTGCTGCACTCCGGCTTCCCCCAGCACACAGGTGATGTGGCTGCCTAGCAGGC
AGAGGCCCTGGTGCAGAAAGTGGCTCCCGGCTTCCAACAGACATCAGGCTCCAAGGAGGCAG
TCGCCAGCCTGCGGAGAGGCCACATCCAGCGGCTGAACCTGCGCTACACTCAGGTGTCCCGCC
AGGAGCTCAGGGACTACGCGTCTTGAACCTGAGTGTGAACCTCTACACAGTCAACGCACCGT
GGCTCTTCTCCCTGCTGTGGTGTGCGGGGTCCCATCCGTACCTCTGACAACTCCCACACCC
TGTCCCAGGTGCCTTCCCCCTCTGGATCATGCCCCCGGACGAGTACTGTCTCATGTGGGTCA
CTGCCGACCTGGTCTCCTTCACCCTCATCGTGGGCATCTTCGTGCTCCAGAAGTGGCGCCTGG
GTGGCATAACGAGCTACAACCCTGAGCAGATCATGCTGAGTGTGCGGTGCGCCGACAGCC
GGGACGTGAGCATCATGAAGGAGAAGCTTATTTTCTCAGAGATCAGCGATGGTGTAGAGGTCT
CCGATGTGCTCTCCGTATGTTGAGACAACAGTTATGACACATATGCCAACAGCACCGCCACCC
CTGTGGGCCCCCGAGGGGGTGGCAGCCACACCAAGACCCTCATAGAGCGGAGTGGGCGT**TAGC**
TGAAGACATGTCTGTCCCACCTGTACCTGACACAGAAGCTGGGGAGCCTAGGAGAGCTGGTGG
AAGTGTGTCTGAACTCGGAGTGTCTGGGAGCGGGCTCCACAGCCTCCTTGTGGGCTCCAGCC
CCTTGTCAGCCGCAGCCTCTCTTGAGGGGGACTCCCTGTCTCCTGAGGCCAGCTGGGCCAGG
ACTCCATCCTTTCAGATGCCCCCTGCAGGCCTGGGGCTCCTTCTGGGAAGTATGGGGCCTAGGG
CTTGGTCCCCCTCTTCTGAGGGCCCTCTCCTGTATCCCGACCTGGAAGCTTTGATGGGTGATGG
GCCATGCCATACCCCCTGTGGCAATGGAGTGTGTGGATGCTCACCTGTGCCATCTGTCTCCT
GTCTGTGCCAGGAGGCACCTGAGTTCTCTGCTGTTATCCTGCCCCAAGGGCCTGGGCCGAGCC
TCTACCTGAAGCAACTCTGCTCTTCTGTGCTCTCAAAGCACAAGGAGGTTGAGCCAGGAG
GAAGCCAGCTGCAATGTGGAGACACGTCCTCCTCCCCAACCCACCTCATGCCACCGCCAACCC
CCTGCCCCAGGAGCGGGCCTGAGCCACGTCCTCCTAGGAGCAGCTGGAGATGGCCAAAAGAGTG
AGCTCAGGACTACTGGATCCCATGCCAGGTGTCCAGCAGACCTCAAGGCAGAAGGGTACCT
AACCCAGGAGTCCACAGACTGATGTGACCTCAGGTTCCACATCAGTGGCCACAGGGCAGGGC
CCACCTGGTAGAAGTGTTCTGGATATGGCCAGGGTGGGTGTGTGGCTAAGTGGGCCTGAACAG
AGGGAACCTAGGGCCCTTGGCCAATGTGATTAAAGCTGCCATCTTGAAA

58/550

FIGURE 58

MVRHQPLQYYEPQLCLSLTGIYGCRWKRYQRSHDDTTPGTAPFLHVGAVAAVTMLSWIVAGQ
FARAERTSSQVTILCTFFTVVFALYLAPLTISSPCIMEKKDLGPKPALIGHRGAPMLAPEHTL
MSFRKALEQKLYGLQADITISLDGVPFLMHDTTLRRTTNVEEEFPELARRPASMLNWTTLQRL
NAGQWFLKTDPFWTASSLSPSDHREAQNQSICSLAELLELAKGNATLLLNLRDPPREHPYRSS
FINVTLEAVLHSGFPQHQMVLPSRQRPLVRKVAPGFQQTSGSKEAVASLRRGHIQRLNLRYT
QVSRQELRDYASWNLSVNLTYVNAPWLFSLWCAGVPSVTSDNSHTLSQVPSPLWIMPPDEYC
LMWVTADLVSEFTLIVGIFVLQKWRLGGIRSYNPEQIMLSAAVRRTSRDVSIMKEKLIFSEISD
GVEVSDVLSVCSDNSYDTYANSTATPVGPRGGGSHTKTLIERSGR

Important features of the protein:**Signal peptide:**

amino acids 1-24

Transmembrane domains:

amino acids 47-61, 77-93, 335-350, 380-399

N-glycosylation sites.

amino acids 182-186, 217-221, 233-237, 255-259, 329-333, 462-466

Tyrosine kinase phosphorylation site.

amino acids 130-139

N-myristoylation sites.

amino acids 21-27, 48-54, 294-300, 404-410, 442-448, 473-479

59/550

FIGURE 59

CCTGAGCAAAACACAGCAGCCCCGAGTGTTCCTCAAGGCCAAAATGCTGAGAACGTCCACTCCTAA
TCTGTGTGGTGGTCTGCATTGCCGGGCCCCCTGGCTCTCTTCTGGCATTCTCTGCCTCTGCCT
CATATTCTTGTAGGCCAGGTGGGCTTGCTGCAGGGACACCCCCAGTGCCTGGATTACGGGCC
CCCTTTCCAGCCCCCTCTGCACCTTGAGTTTGTCTGACTATGAGTCCTTCGGCTGCTGTGA
TCAGCACAAGGACCGCCGCATCGCTGCCCGGTACTGGGACATCATGGAATATTTTGATCTGAA
GAGACATGAGCTGTGTGGAGATTACATTAAAGACATCCTTTGCCAGGAGTGTCTGCCCTACGC
AGCCACCTCTACGACGCCGAAAACACCCAGACGCCTCTCCGGAATCTCCCGGGCCTCTGCTC
TGATTACTGCTCTGCCTTCCATTCTAACTGTCACTCAGCCATTTCCCTGCTGACCAATGACCG
CGGCTTCCAGGAGTCTCATGGAAGGGACGGTACCCGCTTCTGCCACCTCCTGGACCTTCCTGA
CAAGGACTATTGCTTCCCTAATGTCTGAGGAACGACTATCTCAACCGCCACCTGGGCATGGT
GGCCCAAGATCCTCAGGGCTGCCTGCAGCTCTGCCTGAGCGAGGTGGCCAACGGGCTGAGGAA
CCCCGTCTCCATGGTCCATGTCTGGGGACGGCACCCATCGCTTCTTTGTTGCCGAGCAGGTAGG
AGTGGTGTGGGTCTACCTCCCTGATGGGAGTCGCCTGGAGCAACCCTTCCTGGACCTCAAGAA
CATCGTGTGACCACCCCATGGATCGGGGATGAGAGAGGCTTCTTGGGGTTGGCTTTTACCC
CAAATTCGCCACAATCGCAAGTTCTATATTTATTATTCGTGCCTGGACAAGAAGAAGGTAGA
AAAGATCCGAATTAGTGAGATGAAGGTTTCTCGGGCTGATCCTAACAAAGCTGACCTGAAATC
AGAGAGGGTCATCTTGGAGATTGAAGAACCAGCCTCAAACCATAATGGCGGACAACCTTCTTTT
TGGCCTGGATGGCTATATGTACATATTCCTGAGGGACGGGGACAGGCTGGAGATCCCTTTGG
CCTGTTTGGAAATGCTCAGAACAAAAGTTCCCTGCTGGGAAAAGTTTAAAGGATCGATGTGAA
CAGGGCAGGCTCACATGGCAAGCGGTACCGAGTCCCTCGGACAATCCATTTGTTTCTGAGCC
AGGGGCCCCACCCGCCATCTATGCCTATGGGATCAGGAACATGTGGCGTTGTGCTGTGGACCG
AGGGGACCCCATCACGCGCCAGGGCCGAGGCCGATATTCTGTGGGGACGTGGGCCAGAACAG
GTTTGAAGAGGTTGACCTCATTTTGAAAGGTGGAAACTATGGCTGGAGAGCAAAGGAAGGGTT
TGCATGTTATGACAAAAAATTTGTCACAATGCCTCTTTGGATGATGTTCTGCCAATCTATGC
TTATGGCCATGCAGTGGGGAAGTCAGTCACTGGAGGTTATGTCTATCGTGGTTGTGAATCCCC
AAATCTCAATGGCCTGTATATCTTTGGAGACTTCATGAGTGGTTCGACTTATGGCTTTGCAGGA
AGATAGAAAAAAACAAGAAATGGAAGAAGCAGGATCTTTGCCTGGGCAGCACACGTCCTGTGC
CTTCCCAGGGCTGATCAGCACCCATAGCAAGTTCATCATCTCCTTTGCTGAAGATGAAGCAGG
GGAGCTGTATTTCTGGCGACCTCTTACCCAAGTGCCTATGCACCACGTGGATCTATTTACAA
GTTTGTGACCCCTCAAGGCGAGCACCCCCAGGCAAGTGCAATAACAAGCCAGTGGCCGTGAG
AACCAAGAGTAAGCGGATCCCGTTCAGACCACTCGCCAAGACAGTCTTGGACTTGCTAAAGGA
ACAATCAGAGAAAGCTGCTAGAAAATCTTCCAGTGCAACCTTAGCTTCTGGCCCAGCCCAGGG
TTTGTCTGAGAAAGGCTCCTCCAAGAAGCTGGCTTCTCCTACAAGCAGCAAGAATACATTGCG
AGGGCCTGGTACAAAGAAGAAAGCCAGAGTGGGGCCCCACGTCCGCCAGGGCAAGAGGAGGAA
GAGCCTGAAAAGCCACAGTGGCAGGATGAGGCCATCAGCAGAGCAGAAGCGAGCTGGCAGAAG
TCTCCCTTGAACCTATTGGTCAAGGTGGCCGACAGGGTGACGTGAGAGAGGAGAGCCACCTCAT
CAAATGAAAGTCACTGCTGAATAAAGACCTTAGAAGTCTGGGAAGCCAGGGTAGAGGTGGGGC
AGGGCGGTTTTCTCTCCCTGGGAAATCTTGCTGTCTACTGAATAAATAAATGCACCTTCTCT
GTATGCAGTGCTTCTGTGGGAGACCATATCCAGATTGCTGGTGCACCTGGGTTATGGTAAGC
ACTAGTCCATGAGCCTGCTTGGAAATCACACTGGATGTCTCCGTTTTGTCTTGTAATGCCTAC
AACCTGAGGTAATAAATCAACATTTGCTCA

60/550

FIGURE 60

MLRTSTPNLCGGLHCRAPWLSSGILCLCLIFLLGQVGLLQGHQPCLDYGPPFQPPLHLEFCSD
YESFGCCDQHKDRRIAARYWDIMEYFDLKRHELCDYIKDILCQECSPYAAHLYDAENTQTPL
RNLPGLCSDYCSAFHSNCHSAISLLTNDRGLQESHGRDGTRFCHLLDLPDKDYCFPNVLRNDY
LNRHLGMVAQDPQGCLQLCLSEVANGLRNPVSMVHAGDGTFRFFVAEQVGVVWVYLPDGSRL
QPFLDLKNIVLTTPWIGDERGFLGLAFHPKFRHNRKFYIYYSCLDKKKVEKIRISEMKVSRAD
PNKADLKSERVILEIEEPASNHNQGQLLFLGLDGYMYIFTGDGGQAGDPFGLFGNAQNKSSLLG
KVLRIDVNRAGSHGKRYRVPSDNPFVSEPGAHPAIYAYGIRNMWRCAVDRGDPITRQGRGRIF
CGDVGQNRFEEDVLILKGGNYGWRAKEGFACYDKKLCHNASLDDVLPYIYAYGHAVGKSVTGGY
VYRGCEPNLNGLYIFGDFMSGRLMALQEDRKNKKWKKQDLCLGSTTSCAFPGLISTHSKFII
SFAEDEAGELYFLATSYPSAYAPRGSIIYKFVDPSRRAPPGKCKYKPVVVRTKSKRIPFRPLAK
TVLDLLKEQSEKAARKSSSATLASGPAQGLSEKGSSKKLASPTSSKNTLRGPGTKKKARVGPH
VRQGRKRRKSLKSHSGRMRPSAEQKRAGRSLP

Important features of the protein:**Signal peptide:**

amino acids 1-41

Transmembrane domain:

amino acids 17-36

N-glycosylation sites.

amino acids 372-376, 480-484

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 645-649, 699-703

Tyrosine kinase phosphorylation site.

amino acids 81-89

N-myristoylation sites.amino acids 11-17, 37-43, 156-162, 165-171, 357-363, 365-371,
368-374, 408-414, 459-465, 548-554, 557-563**Amidation sites.**

amino acids 391-395, 696-700

Cell attachment sequence.

amino acids 428-431

Leucine zipper pattern.

amino acids 25-47

61/550

FIGURE 61

CTCCATTAAACCACCACCAGCTCCCCAAGCCACCCCTTCAGCCATGAAGTTCCTGCTCCTGGT
CTTGGCAGCCCTCGGATTCCTGACCCAGGTGATCCCAGCCAGTGCAGGTGGGTCAAATGTGT
GAGTAACACCCCAGGATACTGCAGGACATGTTGCCACTGGGGGGAGACAGCATTGTTTCATGTG
CAACGCTTCCAGAAAATGCTGCATCAGCTACTCCTTCCTGCCGAAGCCTGACCTACCACAGCT
CATCGGTAACCACTGGCAATCAAGGAGAAGAAACACACAAAGGAAAGACAAGAAGCAACAAAC
GACCGTAACATCATTAATAACCACTGCTATCGCCTCCACCAACTCAGAGAAATATCATTTCAC
AGTTCCAATTCCTCCTACATTGCTGAGTACTAGCCAAGGCTCCTCTTTATGGGGCAGATATCT
ATAGCCAACCCCAAACCTTCTGTCTTCTATCATTCTGTTCATTCTAGTAACTAATTTGGAG
TTTGTATCTATCTTACGAGAACAAATCATCATGCAGATTTCGTCCACAGGGGATCTGTCAGTTTG
GGTCTCCAAATGAAAAATGTCAAGACAGAATTGGACATGCAAAAGATTGACTGGGAGAACAC
ACCTCTGATGGACAAAGGTGAGACAGAGCAGCCACAGGCAGGGAGAGCCTTCAGACTGCAACG
CTGGCCTGATACGTGTCAAAGGAGAGAGGGATAGAGGAGGATTGAATAGAAGGAGACTAAGAC
TGCAGCTCTAAGAAAGTCTCAGCCAAACAGATGGGGAGGCCCAAAGCAAGGCTTGCCCTCAG
AGGAGCTCACGCAGGGCAGGAATAGCCAGGTTCTCATATCCCAGGGGTTTCAGACTTGGCTGAG
AACAGCCCCTGGAGAACATGGGGTGACTGCTACCATAGGTCTGGAAGTATGAGGCTGTCCACC
AACTATCCCCTTGAAGCAAGTTCTCTTGAAAGGAAATCTAAACAGTGCACCCCCATGGCTGCC
ACGGAGTATAAGGAGGGAGAGAAAGGAGCTGAAAGTCTAGGTTTGGCCAGCTAGGTAGACTGA
CTTGTGAGGTATTTATTTATTCATTTGAGTAACAAAGCAGACAGAATACATAGCCACCATTGG
TAGTACACCCCAAAGCAAGGATGGCATGATGCTGGTGACTCAAACGTGCCTACTCATGGTGT
CAAATTGGCATAATCCTCTTGGAAGCTGTGTGGAAATAAGCACAGAGAAGCAGAACTCTAAT
TGCTTAATCCACTAAACATTACTTCTGGGAATTGGCTCATCATAAATTATCCAAGAGAAAGCA
CAAAGTTATGGGCACAAAGGTTTTCCATATAATATTATTTAAAATGCTGAGAAAATGAAAAA
TCTAAATGGTGAAATATATACTAATGCCATCTATAAATACAAACAAATAGAATGTTTATAGAA
TAATGGAACATAATAACATTATTCAAATTTGCATTTATGCTATAGTTGTCAAATTTGTCTCCT
TATATGATACAAAACCTCATGAAAATTATGACTTTTTTGTGTTGGTTGGAAAGCAGAATTATGCA
TAAATTTCTCTTACAGTTCGATGCCCATTAGTTTTATATAACATTTATTTGACACGTACTGA
CTTCTATCTGAGAAGAACAACCAAAACACTCAGGCCTAAATAATTAAAAACGGTCCTAAAAA
CTAGCAAACCAGATAAGAAAAGATGTAAATGCCCATTCCTAACTTATGTCTTAGACCAAAAT
TAATCTAGATGGTTTTAAATGACAGTGTAAGGTAAGTATTAAAGATTGTGTGGTCAAA
TATTCAATTTAAGAGCAAGGAAATCTTATAAATATAACAATAGAGGCAGAACTCATGTAAGA
ATAAATTGATTAGGTGGTATTAAATATTAAGTTCTTATGTATGTCAAAGATATCATTTTGAA
ATTCATCCATCTTATTGGGTATTGCAGGAGTTCATTCCTTTTTGTTTATAAATACTCTTCCGT
CATATGAATAGTATTCATTTGTATACTGGTTTGTTGATGGACATTTGGGTGTTCCAGTTTA
TGGCTATTACAAATAAAGCTTCTATGAACATTTATGTACA

62/550

FIGURE 62

MKFLLLVLAALGFLTQVIPASAGGSKCVSNTPGYCRTCCHWGETALFMCNASRKCCISYSFLP
KPDLPQLIGNHWQSRRRNTQRKDKKQQTTVTS

Important features of the protein:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 1-22

N-glycosylation site.

amino acids 50-53

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 79-82

N-myristoylation site.

amino acids 23-28

63/550

FIGURE 63

GCGGAGCGCCTGGGAGAGGAGAAGGAGCCGACCTGCCGAGATGGAGGCGACCGGCACCTGGGC
GCTGCTGCTGGCGCTGGCGCTGCTCCTGCTGCTGACGCTGGCGCTGTCCGGGACCAGGGCCCCG
AGGCCACCTGCCCCCGGGCCACGCCGCTACCACTGCTGGGAAACCTCCTGCAGCTACGGCC
CGGGGCGCTGTATTACGGGCTCATGCGGCTGAGTAAGAAGTACGGACCGGTGTTACCATCTA
CCTGGGACCTGGCGGCCTGTGGTGGTCTGGTTGGGCAGGAGGCTGTGCGGGAGGCCCTGGG
AGGTCAGGCTGAGGAGTTCAGCGGCCGGGGAACCGTAGCGATGCTGGAAGGGACTTTTGATGG
CCATGGGGTTTTCTTCTCCAACGGGGAGCGGTGGAGGCAGCTGAGGAAGTTTACCATGCTTGC
TCTGCGGGACCTGGGCATGGGGAAGCGAGAAGGCGAGGAGCTGATCCAGGCGGAGGCCCGGTG
TCTGGTGGAGACATTCCAGGGGACAGAAGGACGCCATTTCGATCCCTCCCTGCTGCTGGCCCA
GGCCACCTCCAACGTAGTCTGCTCCCTCCTCTTTGGCCTCCGCTTCTCCTATGAGGATAAGGA
GTTCCAGGCCGTGGTCCGGGCAGCTGGTGGTACCCTGCTGGGAGTCAGCTCCAGGGGGGTCA
GACCTACGAGATGTTCTCCTGGTTCCTGCGGCCCCCTGCCAGGCCCCACAAGCAGCTCCTCCA
CCACGTACGACACCTTGGCTGCCTTCACAGTCCGGCAGGTGCAGCAGCACCAGGGGAACCTGGA
TGCTTCGGGCCCCGCACGTGACCTTGTGATGCCTTCCTGCTGAAGATGGCACAGGAGGAACA
AAACCCAGGCACAGAATTACCAACAAGAACATGCTGATGACAGTCATTTATTTGCTGTTTGC
TGGGACGATGACGGTCAGCACCACGGTCGGCTATACCCTCCTGCTCCTGATGAAATACCCTCA
TGTCCAAAAGTGGGTACGTGAGGAGCTGAATCGGGAGCTGGGGGCTGGCCAGGCACCAAGCCT
AGGGGACCGTACCAGCCTCCCTTACACCGACGCGGTTCTGCATGAGGCGCAGCGGTGCTGGC
GCTGGTGGCCATGGGAATACCCCGCACCTCATGCGGACCACCCGCTTCCGAGGGTACACCCT
GCCCCAGGGCACGGAGGTCTTCCCCCTCCTTGGCTCCATCCTGCATGACCCCAACATGTTCAA
GCACCCAGAAGAGTTCAACCCAGACCGTTTCTGGATGCAGATGGACGGTTCAGGAAGCATGA
GGCGTTCTTGCCCTTCTCCTTAGGGAAGCGTGTCTGCCTTGGAGAGGGCCTGGCAAAAGCGGA
GCTCTTCTCTTCTTACCACCATCCTACAAGCCTTCTCCCTGGAGAGCCCGTGCCCGCCGGA
CACCTGAGCCTCAAGCCCACCGTCAGTGGCCTTTTCAACATTCCCCAGCCTTCCAGCTGCA
AGTCCGTCCCCTGACCTTCACTCCACCACGCAGACCAGATGAAGGAAGGCAACTTGGAAGTG
GTGGGTGCCCAGGACGGTGCTCCAGCCTCAACAGTGGGCATGGACAGGGTTAATGTCTCCAG
AGTGTAACACTGCAGGCAGCCACATTTACACGCCTGCAGTTGTTTTCCGGAGTCTGTCCACGG
CCCACACGCTCACTTGACTCATGCTGCTAAGATGCACAACCGCACACCCATACACAACACTACAA
GGGCCACAAAGCAACTGCTGGGTAGCTTTCACAGACATAAATATAGTCCATCTGCAATCAC
AAGCACATAGCCAGGTAACCCACCAACTCCCCTGGATCTGCAGCCCACACGTGGGAGTCTGGC
TGTCACCTTCACAAGCCACAGAAACGGCCACACATGTTTACAGCTCACACGCCCTCTCCATT
ATCGAACTTCTCAGTGTCCCTGTCCCTGGTGCCTGGCACAGGGAACAGCATGCCCCCTCCGGG
GTCATGCCACCCAGAGACTGTGCTGTCTATGGCCCCAACTCATGCTCCCTCTCTTGGCTACA
CCACTCTCCCAGCCTGTGACCACCGATGTCCACACACCCCCAACCACTTGTCCACACAGCTAC
CCACGTACAACATCGTCCTGGCTCCCCAGAGTATCTTCCCCTGAGACACGCCGCCCCACAG
AGGCACAGTCCCCAGCCACCTCTGCAACTGCAGCCCTCAGTCACCCCTTTTAAAGCACCTGA
TTCTACCAATGCAAACACATCTGGGTCTGCGATTATGCACAGAGACTTTGGACATACGAGGA
CCCTCAGACCGGAGGAACACCTGCCCAACCCCAACAGTGCTTATGTAACCACGTGGAAAGCG
GCCCCTGCTGCCCTCCACACACACATAACACTCACTGATCTACAGCCCCGTTCGGCGTCA
GAGTCCCCACTAGACCCAGTGGAAGGGGTAGAGACCAAGTAGGGGCCAGTTTCCAATTCACC
CTGTCAGGGAGTGAGCCGATCTGACGTTTCTTGTGACTTAAGGGTCCGGCTTGGGAATTA
GTTTGTCTTCTGGCCTTTAGCCTAAAAA

64/550

FIGURE 64

MEATGTWALLLALALLLLLTALSGTRARGHLPPGPPTPLPLLGNLLQLRPGALYSGLMRLSKK
YGPVFTIYLGWP RPVVVLVGQEAVREALGGQAE EFSGRGT VAMLEGT FDGHGVFFSNGERWRQ
LRKFTMLALRDLGMGKREG EELIQAEARCLVET FQGTEGRPFDP SLLLAQATSNVVC SLLFGL
RFSYEDKEFQAVVRAAGGTLLGVSSQGGQTYEMFSWFLRPLPGPHKQLLHHVSTLAAFTVRQV
QQHQGNLDASGPARDLVDAFLLKMAQEEQNPGTEFTNKNMLMTVIYLLFAGTMTVSTTVGYTL
LLLMKYPHVQKWVREELNRELGAGQAPSLGDRTRLPYTDAVLHEAQRLLALVPMGI PRTL MRT
TRFRGYTLPQGTEVFPLLGSILHDPNIFKHPEEFNPDRFLDADGRFRKHEAFLPFSLGKRVCL
GEGLAKELELFFFTTILQAFSLES PCPPDTLSLKPTVSGLFNIPPAFQLQVRPTDLHSTTQTR

Important features of the protein:**Signal peptide:**

amino acids 1-28

Transmembrane domain:

amino acids 294-313

Glycosaminoglycan attachment site.

amino acids 99-103

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 128-132

N-myristoylation sites.amino acids 51-57, 109-115, 115-121, 188-194, 207-213, 257-263,
284-290, 339-345, 370-376, 444-450**Amidation sites.**

amino acids 140-144, 435-439

Leucine zipper pattern.

amino acids 32-54, 39-61

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 433-443

65/550

FIGURE 65

CGGACGCGTGGGGCCGTATGCGCGGCTCTGTGGAGTGCACCTGGGGTTGGGGGCACTGTGCCC
CCAGCCCCCTGCTCCTTTGGACTCTACTTCTGTTTGCAGCCCCATTTGGCCTGCTGGGGGAGA
AGACCCGCCAGGTGTCTCTGGAGGTCATCCCTAACTGGCTGGGCCCCCTGCAGAACCTGCTTC
ATATACGGGCAGTGGGCACCAATTCCACACTGCACTATGTGTGGAGCAGCCTGGGGCCTCTGG
CAGTGGTAATGGTGGCCACCAACACCCCCACAGCACCTGAGCATCAACTGGAGCCTCCTGC
TATCCCCTGAGCCGATGGGGGCCTGATGGTGCTCCCTAAGGACAGCATTAGTTTTCTTCTG
CCCTTGTTTTTACCAGGCTGCTTGAGTTTGACAGCACCAACGTGTCCGATACGGCAGCAAAGC
CTTTGGGAAGACCATATCCTCCATACTCCTTGGCCGATTTCTCTTGGAACAACATCACTGATT
CATTGGATCCTGCCACCCTGAGTGCCACATTTCAAGGCCACCCCATGAACGACCCTACCAGGA
CTTTTGCCAATGGCAGCCTGGCCTTCAGGGTCCAGGCCTTTTCCAGGTCCAGCCGACCAGCCC
AACCCCTCGCCTCCTGCACACAGCAGACACCTGTGAGCTAGAGGTGGCCCTGATTGGAGCCT
CTCCCCGGGGAAACCGTTCCTGTTTGGGCTGGAGGTAGCCACATTGGGCCAGGGCCCTGACT
GCCCCTCAATGCAGGAGCAGCACTCCATCGACGATGAATATGCACCGGCCGTCTTCCAGTTGG
ACCAGCTACTGTGGGGCTCCCTCCCATCAGGCTTTGCACAGTGGCGACCAGTGGCTTACTCCC
AGAAGCCGGGGGGCCGAGAATCAGCCCTGCCCTGCCAAGCTTCCCCTCTTCATCCTGCCTTAG
CATACTCTCTTCCCAGTCACCCATTGTCCGAGCCTTCTTTGGGTCCCAGAATAACTTCTGTG
CCTTCAATCTGACGTTTCGGGGCTTCCACAGGCCCTGGCTATTGGGACCAACACTACCTCAGCT
GGTCGATGCTCCTGGGTGTGGGCTTCCCTCCAGTGGACGGCTTGTCCCCACTAGTCCTGGGCA
TCATGGCAGTGGCCCTGGGTGCCCCAGGGCTCATGCTGCTAGGGGGCGGCTTGGTTCTGCTGC
TGCACCACAAGAAGTACTCAGAGTACCAGTCCATAAATTAAGGCCCGCTCTCTGGAGGGAAGG
ACATTACTGAACCTGTCTTGCTGTGCCTCGAAACTCTGGAGGTTGGAGCATCAAGTTCCAGCC
GGCCCTTCACTCCCCCATCTTGCTTTTCTGTGGAACCTCAGAGGCCAGCCTCGACTTCCTGG
AGACCCCCAGGTGGGGCTTCCTTCATACTTTGTTGGGGGACTTTGGAGGCGGGCAGGGGACAG
GGCTATTGATAAGGTCCCCTTGGTGTTGCCTTCTTGCATCTCCACACATTTCCCTTGATGGG
ACTTGCAGGCCTAAATGAGAGGCATTCTGACTGGTTGGCTGCCCTGGAAGGCAAGAAAATAGA
TTTATTTTTTTTTCACAGGGAAAAA

66/550

FIGURE 66

MRGSVECTWGWGHCAPSPLLLWTLTLLFAAPFGLLGKTRQVSLEVI PNWLGPLQNLLHIRAVG
TNSTLHYVWSSSLGPLAVVMVATNTPHSTLSINWSLLLSPEPDGGLMVLPKDSIQFSSALVFTR
LLEFDSTNVSDTAAKPLGRPYPPYSLADFSWNNITDSLDPATLSATFQGHMNDPTRTFANGS
LAFRVQAFSRSSRPAQPPRLLHTADTCQLEVALIGASPRGNRSLEFGLEVATLGQGPDCCPSMQE
QHSIDDEYAPAVFQLDQLLWGS LPSGFAQWRPVAYSQKPGGRESALPCQASPLHPALAYSLPQ
SPIVRAFFGSQNNFCFNLTFGASTGPGYWDQHYLSWSMLLGVGFPPVDGLSPLVLGIMAVAL
GAPGLMLLGGGLVLLLHHKKYSEYQSIN

N-glycosylation sites:

amino acids 65-69, 95-99, 134-138, 159-163, 187-191, 230-234,
333-337

cAMP- and cGMP-dependent protein kinase phosphorylation site:

amino acids 397-401

Casein kinase II phosphorylation sites:

amino acids 151-155, 249-253, 255-259

N-myristoylation sites:

amino acids 3-9, 63-69, 235-241, 273-279, 292-298, 324-330

Leucine zipper pattern.

amino acids 371-393

FIGURE 67

[illegible]

68/550

FIGURE 68

MFPLRALWLWVWALLGVAGSCPEPCACVDKYAHQFADCAYPEGLPANVTTLSSLANKI
TVLRRGAFADVTQVTSWLAHNEVRTVEPGALAVLSQLKNLDLSHNFISFPWSDLRNLSALQ
LLKMNHNRGLSPRDALGALPDLRSLRINNRLRTLAPGTFDALSALSHLQLYHNPFHCGCGL
VWLQAWAASTRVSLPEPDSIACASPPALQGVFVYRLPALPCAPPSVHLSAEPPEAPGTPLRA
GLAFVLHCIADGHPTPRLQWQLQIPGGTVVLEPPVLSGEDDGVGAEEGEGEGDGDLLTQTQAAQ
TPTPAPAWPAPPATPRFLALANGSLLVPLLSAKEAGVYTCRAHNELGANSTSIRVAVAATGPP
KHAPGAGGEPDGQAPTSEKSTAKGRGNSVLPSPKPEGKIKGQGLAKVSILGETETETEPEEDTSE
GEEAEDQILADPAEEQRCGNGDPSRYVSNHAFNQSAELKPHVFELGVIALDVAEREARVQLTP
LAARWGPGPGGAGGAPRPGRRPLRLLYLCPAGGGAHVQWSRVEEGVNAYWFRGLRPGTNYSVC
LALAGEACHVQVVFSTKKELPSLLVIVAVSVFLLVLATVPLLGAACCHLLAKHPGKPYRLILR
PQAPDPMEKRIAADFPRASYLESEKSYPAGEAGGEEPEDVQGEGLDEDAEQGDPSGDLQRE
ESLAACSLVESQSKANQEEFEAGSEYSDRLPLGAEAVNIAQEINGNYRQTAG

Important features of the protein:**Signal peptide:**

amino acids 1-19

Transmembrane domain:

amino acids 587-610

N-glycosylation sites.

amino acids 52-55, 121-124, 337-340, 364-367, 474-477, 563-566

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 397-400

Casein kinase II phosphorylation sites.amino acids 19-23, 202-205, 289-292, 246-249, 411-414, 431-434,
433-436, 440-443, 544-547, 583-586, 650-653, 700-703**N-myristoylation sites.**amino acids 15-20, 48-53, 165-170, 296-301, 351-356, 362-367,
390-395, 419-424, 514-519, 536-541, 557-562, 561-566, 610-615,
661-666, 716-721**Amidation site.**

amino acids 522-525

Prokaryotic membrane lipoprotein lipid attachment sites.

amino acids 10-20, 603-613

69/550

FIGURE 69

GGCGGCGGGAGCAGCGAAGGGGGCGGCAGGGATCCTCCAGGCTGCCGGCTGGGAAGGCGTGGG
CGACCCGGTGTGTGGCGCGCCAGAGCCCCGCGTTTCAGCCCTAGGGAAGGAAGCCAGTTGAG
GGAAGTTCTCCATGAATGTACGTCACAATGATGATGACCGACCAAATCCCTCTGGAACCTGCCA
CCATTGCTGAACGGAGAGGTAGCCATGATGCCCCACTTGGTGAATGGAGATGCAGCTCAGCAT
GTTATTCTCGTTCAAGTTAATCCAGGTGAGACTTTTACAATAAGAGCAGAGGATGGAACACTT
CAGTGCATTCAAGGACCTGCTGAAGTTCCCATGATGTCACCCAATGGATCCATTCTCTCCCAT
CATGTGCCTCCAGGTTATATCTCACAGGTGATTGAAGATAGTACTGGAGTCCGCCGGGTGGTG
GTCACACCCCAGTCTCCTGAGTGTTATCCCCCAAGCTACCCCTCAGCCATGTCTCCAACCCAT
CATCTCCCTCCCTATCTGACTCACCATCCACATTTTATTCACTACACACGGCTTACTAC
CCACCTGTTACCGGACCTGGAGATATGCCGCCCTCAGTTTTTTCCCCAGCATCATCTTCCCCAC
ACAATATATGGTGAGCAAGAAATTATACCATTTTATGGAATGTCAAGCTACATCACCCGAGAA
GACAGTACAGCAAGCCTCCGCACAAAAAACTGAAAGACCGCCAGATCGATCGCCAGAACC
CTCAACAGCCCTCCTTCTTCTATCTACAAAAGCAGCTGCACAACAGTATACAATGGCTATGGG
AAGGGCCATAGTGGTGAAGTGGCGGAGGCGGCAGCGGTAGTGGTCCCGGAATTAAGAAAA
GAGGACGAGCAAGAAGCAGCCCAAAGTCGAATGATTGAGACTTGCAAGAATATGAGTTGAA
GTAAAGAGGGTGCAAGACATTCTTTCCGGAATAGAGAAACCACAGGTTTCTAATATTCAGGCA
AGAGCAGTTGTGTTGTCCTGGGCTCCCCCTGTTGGACTTTTCTGTGGACCCACAGTGGTCTT
TCCTTCCCCTACAGTTACGAGGTGGCCTTATCAGACAAAGGACGAGATGGAAAATACAAGATA
ATTTACAGTGGAGAAGAATTAGAATGTAACCTGAAAGATCTTAGACCAGCAACAGATTATCAT
GTGAGGGTGTATGCCATGTACAATTCCGTAAAGGGATCCTGCTCCGAGCCTGTTAGCTTCACC
ACCCACAGCTGTGCACCCGAGTGTCTTTCCCCCTAAGCTGGCACATAGGAGCAAAGTTCA
CTAACCTGCAGTGAAGGCACCAATTGACAACGGTTCAAAAATCACCAACTACCTTTTAGAG
TGGGATGAGGGAAAAAGAAATAGTGGTTTCAGACAGTGCTTCTTCGGGAGCCAGAAGCACTGC
AAGTTGACAAAGCTTTGTCCGGCAATGGGGTACACATTCAGGCTGGCCGCTCGAAACGACATT
GGCACCAGTGGTTATAGCCAAGAGGTGGTGTGCTACACATTAGGAAATATCCCTCAGATGCCT
TCTGCACCAAGGCTGGTTTCGAGCTGGCATCACATGGGTACGTTGCAGTGGAGTAAGCCAGAA
GGCTGTTACCCGAGGAAGTGATCACCTACACCTTGGAATTCAGGAGGATGAAAATGATAAC
CTTTTCCACCCAAAATACACTGGAGAGGATTTAACCTGTACTGTGAAAAATCTCAAAAGAAAGC
ACACAGTATAAATTCAGGCTGACTGCTTCTAATACGGAAGGAAAAAGCTGTCCAAGCGAAGTT
CTTGTGTTGTACGACGAGTCCTGACAGGCCTGGACCTCCTACCAGACCGCTTGTCAAAGGCCCA
GTTACATCTCATGGCTTTAGTGTCAAATGGGATCCCCCTAAGGACAATGGTGGTTCAGAAATC
CTCAAGTACTTGCTAGAGATTACTGATGGAAATCTGAAGGTGAAGTTTTTGGCAATTGTTTT
ATTCAAATCCAATAGCAAGCTCTGTTTTCTAATATAGTAAATGCTTTTATAGTAATAGTGAGT
AATCATTAATTCTAAAGATAGAATTATTATTACAATAAACAACTTTAGTCACATATTGGCAG
TTTTTCTATTTCAAACACAGCACCAGAGATCAGAGTCTACTTGAACTTACATTTGTGTTATT
TAACAATTTTTCTGTATCTTTTTCATTGGTGTGTTTTGTTTTGTTTATCTTTTGTGTTTTCT
TTGGTTTGGTTTGTGTTTTGTTTTGTTTTGAGATACGATCTCTGTACACAGGCTGGAGGGC
AGTGGCACAGACATGGCCCATTCAGTCTCAGACTCCTGGGCTTAAGTGACTCTTCTGCCACA
GAAGATGAGGAAGAATACATTTTTCATAGTGATGGGGTCTCACTATGTTATCTAGGCTGGTCT
CAAACCTCTGGCCTCAAGCAACCCTCCACCTTGCCCTCCCAAAGTGCTGGGACTATAGACATG
AATCACCACACTCAGCTTCCATGTCTTTTATGAAGTGGGTTCCCTAATTAATCAGATAAATT
TGGTATTTTCATCTCCTAAGTGGCATATGTTTTCTGGAAATCTTATAAGCAGCCGAGAGTG
GTGGCTCACGCTGTAGTCCCAGCACTTTGGGAGGCTGAGGTGGGTGGTCAGGAGATCAAGACC
ATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGTGTGGTG
GCAGGCACCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGAAAGATTGCTTGAACCCAGCAG
GCGGAGGTTGCAGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGTGACAGAGTGAGACTC
TGTCTCAAAAAAAAAAAAA

70/550

FIGURE 70

MMMTDQIPLELPPLLNGEVAMMPHLVNGDAAQHVILVQVNPGETFTIRAEDGTLQCIQGPAEV
PMMSPNGSIPPIHVPPGYISQVIEDSTGVRVVPVTPQSPECYPPSYPSAMSPTHLPPYLTHH
PHFIHNSHTAYYPPVTGPGDMPPQFFPQHHLPHHTIYGEQEIIIPFYGMSSYITREDQYSKPPHK
KLKDRQIDRQNRNLNSPPSSIIYKSSCTTVYNGYGKGHSGGSGGGSGSGPGIKKTERRARSSPK
SNDSDLQEYELEVKRVQDILSGIEKPQVSNIQARAVVLSWAPPVGLSCGPHSGLSFPYSYEVA
LSDKGRDGKYKIIYSGEELEC NLKDLRPATDYHVRVYAMNSVKGSCSEPVSFTHTHSCAPECP
FPPKLAHRSSSLTLQWKAPIDNGSKITNYLLEWDEGKRNSGFRQCFFGSQKHCKLTKLCPAM
GYTFRLAARNDIGTSGYSQEVVCYTLGNIPQMPSAPRLVRAGITWVTLQWSKPEGCSPEEVIT
YTLEIQEDENDNLFHPKYTGEDLTCTVKNLKRSTQYKFRLTASNTEGKSCPSEVLVCTTSPDR
PGPPTPLVKGPVTSHGFSVKWDPPKDNNGSEILKYLLEITDGNSEGEVFGNCFIQIQ

Important features of the protein:**N-glycosylation sites.**

amino acids 69-73, 254-258, 401-405

Glycosaminoglycan attachment sites.

amino acids 229-233, 234-238, 236-240

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 416-420, 535-539

Tyrosine kinase phosphorylation site.

amino acids 319-326

N-myristoylation sites.amino acids 52-58, 227-233, 228-234, 230-236, 231-237, 232-238,
235-241, 239-245, 402-408, 610-616**Amidation site.**

amino acids 414-418

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 290-301

ATP/GTP-binding site motif A (P-loop).

amino acids 546-554

CUB domain proteins profile.

amino acids 294-301

71/550

FIGURE 71

AAGTCATTCACTGGATGTGATCTTGGCTCACAGGGGACGATGCTCAAGCTCTTCCTGGCTCCTTCTCAGCCTTGTT
GCTGTAAGTCTGCTCAGTCCACCATTGAGGAACAGGCCAAGACATTTTGGACAAGTTTAAACCACGAAGCCGAA
GACCTGTTCTATCAAAGTTCACTTGCTTCTTGGAAATTATAACACCAATATTACTGAAGAGAATGTCCAAAACATG
AATAATGCTGGGGACAAATGGTCTGCCCTTTTAAAGGAACAGTCCACACTTGCCCAAATGTATCCACTACAAGAA
ATTGAGAATCTCAGTCAAGCTTCACTGCAGGCTCTTCAAGAAAATGGGTCTTCACTGCTCTCAGAAGACAAG
AGCAAACGGTTGAACACAATTCTAATAACAATGAGCACCCTACAGTACTGGAAAAGTTTGTAAACCCAGATAAT
CCACAAGAATGCTTATTACTTGAACCAGGTTTGAATGAAATAATGGCAAACAGTTTAGACTACAATGAGAGGCTC
TGGGCTTGGGAAAAGCTGGAGATCTGAGTCTGGCAAGCAGCTGAGGCCATTATATGAAGAGTATGTGGTCTTGAAA
AATGAGATGGCAAGAGCAAATCATTATGAGGACTATGGGATTATTGGAGAGGAGACTATGAAGTAAATGGGGTA
GATGGCTATGACTACAGCCGCGGCCAGTTGATTGAAGATGTGGAACATACCTTTGAAGAGATTAAACCATTATAT
GAACATCTTCATGCCATGTGAGGGCAAAGTTGATGAATGCCATATCCTTCTATATCAGTCCAATTGGATGCCCTC
CCTGCTCATTGCTTGGTGATATGTGGGGTAGATTTTGGACAAATCTGTACTCTTTGACAGTTCCCTTTGGACAG
AAACAAAACATAGATGTTACTGATGCAATGGTGGACAGGCCCTGGGATGCACAGAGAATATTCAAGGAGGCCGAG
AAGTCTTTGTATCTGTTGGTCTTCTTAATATGACTCAAGGATTCTGGGAAAATCCATGCTAACGGACCCAGGA
AATGTTCAAGAAAGCAGTCTGCCATCCACAGCTTGGGACCTGGGGAAGGGCGACTTCAGGATCCTTATGTGCACA
AAGGTGACAATGGACGACTTCTGACAGCTCATCATGAGATGGGGCATATCCAGTATGATATGGCATATGCTGCA
CAACCTTTTCTGCTAAGAAATGGAGCTAATGAAGGATTCCATGAAGCTGTTGGGGAATCATGTCACTTTCTGCA
GCCACACCTAAGCATTTAAATCCATTGGTCTTCTGTCAACCGATTTTCAAGAAGACAATGAAACAGAAATAAAC
TTCTGCTCAACAAGCACTCACGATTGTTGGGACTCTGCCATTTACTTACATGTTAGAGAAGTGGAGGTGGATG
GTCTTTAAAGGGGAAATTTCCCAAAGACCAGTGGATGAAAAGTGGTGGGAGATGAAGCGAGAGATAGTTGGGGTG
GTGGAACCTGTGCCCCATGATGAAACATACTGTGACCCCGCATCTCTGTTCCATGTTTCTGATGATTACTCATTC
ATTCGATATTACACAAGGACCCTTTACCAATTCCAGTTTCAAGAAGCACTTTGTCAAGCAGCTAAACATGAAGGC
CCTCTGCACAAATGTGACATCTCAAACTCTACAGAAGCTGGACAGAACTGTTGTAAGAAATACCTCAAATGTT
GAACCTCTCCTAGTATTCACTATTACTCATTTCCATGCCTAGGTTTGTATTGATTCTTTGTTCTAAAAAGAAA
ATTTTATGGCCTCAAATGTCTCATTTACAAACCAACATTTAATTTGTGGTCAGACAGGAACCTAGACCATAC
AACAATTGGGTGGGCCACCTCTTTTCTCCCTATCATACTACAGCCCTCTTCTCCTGGTAATTGGAAGGAAAGAG
CGGTTAGGGTGGAAATATATCTGTTAATATGCATTCTTTTCTTATCTGCCAGAAGCAAATTTAGCCAAGTCAAAG
AGAAGAAACCATGATGATCATAGTATAATATATGTACATCTGGAACCCCTCAAAGGCCCTGAACCCCTTTT
TGTGTAGCAATATGCTGAGGCTTGGAAAATCAGAACCCCTGGACCCTAGCATTGGAAAATGTTGTAGGAGCAAGAA
CATGAATGTAAGGCCACTGCTCACTACTTTGAGCCCTTATTTACCTGGCTGAAAGACCAGAACAGAATTTCTTT
TGTGGGATGGAGTACCGACTGGAGTCCATATGCAGACCCAAAGCATCAAAGTGAGGATAAGCCTAAATCAGCTC
TTGGAGATAAAGCATATGAATGGAACGACAATGAAATGTACCTGTTCCGATCATCTGTTGCATATGCTATGAGGC
AGTACTTTTTTAAAGTAAAAAATCAGATGATTCTTTTGGGGAGGAGGATGTGCGAGTGGCTAATTTGAAACCAA
GAATCTCCTTTAATTTCTTTGTCACTGCACCTAAAAATGTGTCTGATATCATTCTAGAACTGAAGTTGAAAAGG
CCATCAGGATGTCCCGGAGCCGATCAATGATGCTTTCCGTCTGAATGACAACAGCCTAGAGTTTCTGGGGATAC
AGCCAACACTTGGACCTCCTAACCGACCCCTGTTTCCATATGGCTGATGTTTTTGGAGTTGTGATGGGAGTGA
TAGTGGTTGGCATTGTCACTCTGATCTTCACTGGGATCAGAGATCGGAAGAAGAAAAATAAGCAAGAAGTGGAG
AAAATCCTTTATGCCTCCATCGATATTAGCAAAGGAGAAAAATAATCCAGGATTCCAAAACACTGATGATGTTTCA
CCTCCTTTTAGAAAAATCTATGTTTTCTCTTGGAGGTGATTTTGTGTATGTAAATGTTAATTTCTATGTTATAG
AAAATATAAGATGATAAAGATATCATTAAATGTCAAACCTATGACTCTGTTCAAGAAAAAATTTGTCCAAAGACA
ACATGGCCAAGGAGAGAGCATCTTCATTGACATTGCTTTCACTATTATTTCTGTCTCTGGATTTGACTTCTGTT
CTGTTTCTTAATAAGGATTTTGTATTAGAGTATATTAGGAAAAGTGTGATTTTGGTCTCACAGGCTGTTCAAGGA
TAATCTAAATGTAATGTCTGTTGAATTTCTGAAGTTGAAAACAAGGATATATCATTGGAGCAAGTGTGGATCT
TGTATGGAATATGGATGGATCACTTGTAAAGACAGTGCCTGGGAAGTGGTGTAGCTGCAAGGATTGAGAATGGCA
TGCATTAGCTCACTTTTCAATTAATCCATTGTCAAGGATGACATGCTTTCTTCAAGTAAGTCAAGTCAAGTACTA
TGGTGATTTGCCATAGTGTGTTGGAATCGATCATGCTTTCTTCAAGGTGACAGGTCTAAAGAGAGAAGAAATC
CAGGGAACAGGTAGAGGACATTGCTTTTCACTTCCAAGGTGCTTGTCAACATCTCCCTGACAACACAAAACATA
GAGCCAGGGGCTCCGTGAACCTCCCGAGAGCATGCTGATAGAAAACCTCATTCTACTGTTCTCTAAGTGTGGAGT
GAATGGAATTTCAACTGTATGTTCAACCTCTGAAGTGGGTACCCAGTCTCTTAATCTTTTGTATTGCTCACA
GTGTTTGAGCAGTGTGAGCACAAGCAGACACTCAATAAATGCTAGATTTACAAA

72/550

FIGURE 72

MSSSSWLLLSLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQNMN
NAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQNGSSVLSEDKSKRLNTILNTMSTI
YSTGKVCNPDNPQECLLLEPGLNEIMANSLDYNERLWAWESWRSEVGKQLRPLYEEYVVLKNE
MARANHIEDYGDYWRGDYEVNGVDGYDYSRQGLIEDVEHTFEEIKPLYEHLHAYVRAKLMNAY
PSYISPIGCLPAHLLGDMWGRFWTNLYSLTVFPFGQKPNIDVTDAMVDQAWDAQRIFKEAEKFF
VSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKGDFRILMCTKVTMDDFLTAHHEMGGH
IQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSIGLLSPDFQEDNETEINFLK
QALTIVGTLFPFTYMLEKWRWMVFKGEIPKDQWMKKWEMKREIVGVPEVPVPHDETYCDPASLF
HVSDDYSFIRYYTRTLYQFQFQEQEALCQAAKHEGPLHKCDISNSTEAGQKLL

Important features of the protein:**Signal peptide:**

amino acids 1-17

N-glycosylation sites.

amino acids 53-57, 90-94, 103-107, 322-326, 432-438, 546-550

N-myristoylation sites.

amino acids 260-266, 286-292, 395-401

Cell attachment sequence.

amino acids 204-207

Neutral zinc metallopeptidases, zinc-binding region signature.

amino acids 371-381

73/550

FIGURE 73

CCCACGCGTCCGAGCGGGGTGGACAAGTGGCGTGTGTGCTGCGACCCCGAGGGAAGATGAACG
GGACGCGGAACCTGGTGTACCCTGGTGGACGTGCACCCAGAGGACCAGGCGGCGGGCAGGA
AGACCTATGCCATGGTGTCCAGCCACTCAGCTGGTCATTCTCTGGCTTCAGAACTGGTGGAGT
CCCATGATGGACATGAGGAGATCATTAAAGGTGTACTTGAAGGGGAGGTCTGGAGACAAGATGA
TTCACGAGAAGAATATTAACCAGCTGAAGAGTGAGGTCCAGTACATCCAGGAGGCCAGGAACT
GCCTACAGAAGCTCCGGGAGGATATAAGTAGCAAGCTTGACAGGAACCTAGGAGATTCTCTCC
ATCGACAGGAGATACAGGTGGTGTAGAAAAGCCAAATGGCTTTAGTCAGAGTCCACAGCCC
TGTACAGCAGCCCACCTGAGGTGGACACCTGTATAAATGAGGATGTTGAGAGCTTGAGGAAGA
CGGTGCAGGACTTGCTGGCCAAGCTTCAGGAGGCCAAGCGGCAACACCAGTCAGACTGTGTGG
CTTTTGAGGTCACACTCAGCCGGTACCAGAGGGAAGCAGAACAAAGTAATGTGGCCCTTCAGA
GAGAGGAGGACAGATGTCCAGAGTGAATTGGAGAATGTCTCTGGGGGAATGAAGTTCCTTCCACA
AACACAGCTCAGTTCTTAGCAACAACTGTTTGTCTTTCTACTTGCTCCATCTGCAGCCTACG
CTGCCCTGGCCTCCTGCAGACAGATAGTGGGGTTACCTGGCAAGGCCTGGTGAGAGCCAGTGA
ACCTAAGCTTTGACTGGGTGGCCTTGCTTTCTGGGGAGGAGGGAATGTACATTCAGGGAGTA
GCCTTTTGCGGAAAAATTCTCTAGGGCTACAGACAGTCATGTGTGACTTCTCTCTGCTGTGAA
AACTCCCAGAGTCTCTTTAGGGATTTTCCCTAAGGTGTACCACCAGGCACACCTCAGTCTTCT
TGACCCAGAGCCTGAAAACGTGTTTCACTGGGTCCACCAGTCCCAGCAAAATCCTCTTTGTA
TTTATTTTGCTAAGTTATTGGTGGTTTTGCTTACATCTCATGATTGATATAATACCAAAGTTC
TATAGCCTTCTCTTGCAGTATTTGGATTTGCTTGAAACCGGGAAAACTGTTCCCATTAGGCTT
GTTAATGTCAGAGTGACACTATTATGAATCTTTCTCTCCCTTTCTCTGCCTGTTTCTTCTCT
CTTTCTCCTTCAAACCTTGCTCTGCAGCTAAGGAAGGTGAGTCTACTTTCCCTGAGGCTTTGGG
GTCAGAGTATATGTTGTTTGGAGAAAGAGGGCAATCAGGACTCTTCTGGGACCCAGATGAGTT
CTTCACTAGCCCTTCTGAACCCCTTGCTCCATAATTGGTCTTTTATCCTGGCTCTGAATGACC
CTGCAGGTCATCATGGTTTTCTTTTTTATTGTTTTTTTTTTTTCTGAGACAGAGTCTCACT
CTGTACCCAGGCTGGAGTGCAGTGGCGGATCTCAGCTCACTGCAACCTCTGCCTCCCGGAT
TTAAGCGATTCTTCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGTGTGCCACCACGCCTG
GCTGATTTTTGTATTTTTAGTAGAGATGGGGTTTTACCATACTGGCTAGGCTGGTCTCGAATT
CCTGACCTCAGGTGATCCACCCACCTCGGCTTCCCAAAGTGCTAGGATTATAGGCTTGAGCTA
CTGCGCCCCGGCCCATGGTGTTTTTCTTTAGGGCTCTTCTACAGCCTTGAGAAGTAGATAGGC
ATCAGAGTATGGTACTATAGGAATCAGAAAAATTCAAAACAAATGTGGATTAAGTGTTTAGGC
TCTATGTGGCTCACGCAGCCAGAATCCTTAAGTCTGTGTGTTTCTGTGTCTCAAGACTGGGCT
CACATTCTGGCTTTGTCCATAACAATGCTCTGGGATTTTCAAGGAGTTCCTCATTTGTAAAT
GAGGGGGTCAGAGCAGGTGATATCCATGTTTCTTCCCTTTCTGATATTGTTGTCTGTGGCATA
TTCTTTGTATGGCGAATTTAATAAATTATATTAATGTGTCA

74/550

FIGURE 74

MNGTRNWCTLVDVHPEDQAAAGRKTYAMVSSHSAGHSLASELVESHGHEEIIKVYLKGRSGD
KMIHEKNINQLKSEVQYIQEARNCLQKLREDISSKLDRLGDSLHRQEIQVVLEKPNGFSQSP
TALYSSPPEVDTCINEDVESLRKTVQDLLAKLQEAKRQHQSDCVAFEVTLSTRYQREAEQSNVA
LQREEDRCPE

Important features of the protein:

Signal peptide:

amino acids 1-39

N-glycosylation site.

amino acids 2-6

Amidation site.

amino acids 21-25

75/550

FIGURE 75

GCTTGCACACATGGCTCCGGAGGCTCCGGTTGCCCATCCGAGCCCCTGCCAGGCTCTAACGTTCCCAACTGACAA
CACCAGTAACATAATATAGGAGCAGATGGTGGGGACGGGCTGTGCGAGCGGCTCCTTTGCAGAGGCTCCGGACT
GCAGATAAGGCTCAGGCCCTTTTGTGAGAAGCAGACCAGCCTGGGGGCTGGCGGCAGGACACCTGTGTCTGCATG
CTGAAGAAGATGGGTGAGGCCGTGGCCAGAGTAGCAAGGAAGGTCAACGAGACGGTGGAGAGCGGCTCTGACACT
CTGGACCTGGCCGAGTGCAAGCTGGTCTCCTTTCCCATTTGGCATCTACAAGGTCCTGCGGAATGTCTCTGGCCAG
ATCCACCTCATACCCTGGCTAACACAGAGCTTAAGTCCCTCACCAGCAAGTTCATGACCACATTCAGTCAGCTC
CGAGAGCTCCACCTGGAGGGGAACTTCCTACACCGCCTCCCAGCGAGGTCAGTGCCCTGCAGCACCTCAAGGCC
ATTGACCTGTCCCGGAACAGTTCCAGGACTTCCCTGAGCAGCTTACCGCCTGCCGGCGCTGGAGACCATCAAC
CTGGAGGAGAACGAGATCGTAGATGTGCCCGTGGAGAAGCTGGCCGCCATGCCAGCCTTGCGCAGCATCAACCTC
CGCTTCAACCCACTCAACGCCGAGGTGCGCGTGATCGCCCCGCCGCTCATCAAGTTTGACATGCTCATGTCTCCG
GAAGGCGCAAGAGCCCCCTACCTTAGGCCACCCTCCTCATGCCACCCAGCAAGGGACAGAGGCCACAGGCCTG
GAACCTTGAAGGGAGGGAGGCCCATGGGAGGCCAAGCCTGGGGGCTGGGGGCGGGTGGGCCGAGCAGCAGTGG
TGGGTGGGGTGACGCTGGTCTGGATAGATAGCTTACAGCAGTAGTGGGCTCTGGAATGCCAAGGGAAGAGGCCAA
GGTGGGGCTGCAGCCTGGACTCGGCACTCACAGCTGCTGTGCAAACTCAGGCAGATCTCCTGCCCTCTCTGAGC
CTTGTCACTTGAAGGAAACAGGACCCTTTCCCTCCTTTGGGCTCCCTGGAGGTTTTTAAGCAGTACGTGCCTCCA
AGTTACCTCCAGATCAGCAGGCACAGGTGGGCATTGCCAGGTATTTTCTGAGCCCCTGCGGGTTTGAGGCCTTGT
TTTTAGTGCTGAGAGCCAGTTGCTGCCCTGAGAAGAGAAGACAACCTCCATCTATTTATTGCTTCTTGAGAACTG
ACCTGGATGCGGCCCTCTGCAGGGCCCAGTCTTCACTCCTGTGGTCCCTGGACTGGTGGGAACCTGAAGTAGGAG
TCCTGGGAGAGCTGTGGTGGGAATATGGGCTGGCACTGCTGCAGGGCAAGAACATTATGTAGGAGCCCGAGGAC
CANCANGCTGGGAATGGGAGCAAGTCACGTGAGCTCTGTCATTCCCCACAGTTAACAATTTGGCGGGGTGGGAA
GTCCTGAGTGCTCCGTCCCTCTAGCATCACTCCTGAGCTGCGGGAGAGGTGGCCCCAGAGAACAGCAGAGTCAGTT
ACACCTGCAGCTCTTGTCTAAAGTGATTAGATGGCCACCCTCACCAGTCCAGTCCAGCAGCAGCCTGGCTGCC
TTGTCAATGGCCCTCCTGGGGGCAAGGCGATGTGGACACCGGATTTGTAGCCAGCCAGCTCCAGGCCAACGCC
CAAAGCCCTGATGACCTGGTTCTTCTGAGGCCCTCAACCTGGCATCTTAGGGTATGGTCAGGCAACAGGGTGACC
AGCTGTCTGTTTCCAGGACATGGAACCTTCAATGCTAAAAGTGGGACATTACCCAGCAAGTGGGGATGGTTG
GTCCCCTACAGGAGAGGGCCTGGGGCTCTTGCTTCCCGAGAACGCCCTGTGGCTTGAAGAACCTTGACTGCTTGG
TCCTCAGGTATCTACCTCCCACCTTCTCCTCATCTGTGGAGCAAGCCAACTCAGTGCCCCAGACCCACCTGATC
TGCATCTTTGTTGCTCCAGAGACACCTGAGGCCCCAGAGCTTGAGGCAAGCCAGGCCGTCCAAATCCTGTGTG
CCGTGGACGAGTGGCCACTTTACTACTCCTAAGGCTAAGATGTTGAGAGCTCAGACCACTGCTCAGAGCAGTAAT
CCCTGCTCAGAAATGCTCCAGTTCCTCGTCCCTGCCAGGTCTCTTGTCTCTTGGGAAGGAACTGATAGGTCCG
GCCATTGTTGGGCCATCACTGAGCGCTCAGTATCTCAAGAGACTCTGTTCAATTCTGCTCGTATCCCAAGGCCCTGG
TTGGTCAAACTCTGGGCAAGGGTTTTTCAGGATGAGGAGGTCAAGACAGGATGTCCAGAGCTACCGAGTTCATCT
GTGGGTGTTGGGGCAAGTGGGGGCTGAAGTCTGTGCAGGCTGCGCTGGCCCCACCTGCCTTGTGCCCTGGAGT
GGGGTTTTCTCCTTGTGGAAGAGAGGCATCCTTCTCTGATGTGCACAAACACAATGTATGACCAGAGCCTTGCAA
CTCAAAGTGTGGTCTGTGGACCAGCAGCGGCAGTGACACCTGGGAGCTTGTAGGAATGCAGAGCTTAGGCCTCA
CCCTATACCTCCCGACTCAGACCCTGCATTTTAGCAAGACCCCAAGCTGATTCTATAAGCACTTTAGAGTTTGA
GAAGCAAGGACCTAGGCTGGGGATGTCCTCCGAGCAGAGGGTGAAGTTTCTCTCAGTTCTCTCCCTGCCACTTCC
AGGGATCTGAGCCTGTGTTGAGCCTCCTCCCTAACCCAGCCCTGGGAGACACTTGGCCTGTTAGATTGTTCCAGAG
TCTGCATGGCACTCCTGAAGAAGGGAGTGTGACCTGCAGTCACCAGGAGATGAGGGTTAGGTGTGCCAGCCCTC
CAGACCCGGCCTTTCTGGTTAACCCCTGCATGCCAAGCTGCCTGCTGCCCCAGGTCTCACCCTCAGGCCTTTGAA
GGGGCAGCTTCTGGAAGTTGTTTCTCCTCTGCTTGGAGAGTTTGCCCTTGTCTGTCTTGGAAAGTGTGGGCAGC
CACAGATGCCCCAAATCAGAGCTCACAGTGAAGTGAAGCCCTAAGCTTCACTGCTGCAATAAAGAATGCATTGGTT
TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

76/550

FIGURE 76

MLKKMGEAVARVARKVNETVESGSDTLDLAECKLVSFPIGIYKVLNRNVSGQIHLITLANNELK
SLTSKFMTTFSQLRELHLEGNFLHRLPSEVSALQHLKAIDLSRNQFQDFPEQLTALPALETIN
LEENEIVDVPVEKLAAMPALRSINLRFNPLNAEVRVIAPPLIKFDMLMSPEGARAPLP

Important features of the protein:

N-glycosylation sites.

amino acids 17-21, 47-51

77/550

FIGURE 77

CACCAACAAGCAATCGTTTCATGAGAAAGCCGTGCACCCGCTGCAGTTGGGCCATGTGGTCCGCATCGTATTCCAC
TAGGTCCCCATTGTACACCAAGTACTGTCCCGGCGTCTCCAGCAGATGCCTGCAGCCTTCCACCTTCTCAAGCAG
GGTGGTGTGAGTGCCTGCTTTCTTCTCGCTCGCCTGGACCGGAGCCGTGCGGGGAGGCACCCCGGGGGTGGAGAA
AAAGCCGGCCTGGCCTCGGAGGTGTCTCGGCCCGCCGCCCCACCGACTCCCTCCTCCCTCCAGAGGCGGGGCG
GGCTCCGGCGGCAGCAGCGGCAGCAGCAACGTAAGCGGGATGCTCTCCAGGCTGCTTTTCTGCTCGGTGAGCAA
ATGGCTGAGCTGGTACATCTCGCTCTCCAGGTAGGAGATCTCGCGGGCGCTCTCTATGAACTGCCGGTAGTTCTG
GTAGACGTTGCGCTTCAGGTTCTGCGCCGTCTCTCCGCCAGCGCCTGGATGCGCTGCCGGTGTCTCTGGAGGTC
CCGGTCCCCATCCGACTGCTGCGAGAGCTGCTTACGTACAGCCGCGCCTCAAAACCCCTGACTCCAGCTGCCG
ACGAGGCGGGCTCGCCCCACTGTCCGACATCGCCATCGCCATTTCTCTCGGGTCTCACGCACTCACTGTCACTA
TCGGCGCCGAGCCGCGCGGCTGTCTAGACCCACCCAAGGCCAACCGAGCTCCTGGGCTGAGGAAGCAGGAATG
GGAACGAGACGAGTACGCTGCGCGGGTCTGAGCGTCAGACACTGCGCTGCGCAAGTGGCGGAGCGCAGACA
TTGCGCCTGCGCAGCAATGCCATCGGTTAAAGCGCATGCGCAAGATGAGCTATTGCGGAAGTGAGGGGAGGGAGA
GGCCGAGAGAAATTTCCGTTACTGCGCATGAACCGAGCGTGACGTTGAGGTTTGAATAACCGCAAAGAGTAAAG
GCTGAAACTAGCTTCTGAAAGCTTCGTAGGGCCGAGCCCTGTGAGCCAGGTTCTGCGCCCACTAGGAGGTGT
CATGCTGACTGCTTTTTTTAAAGCCCTAGAATCCTTGGCTTGGCGTTTGGGGTAAGCTCCGTTCTCGTTCTCAA
GCGCGTTTCCGCGAACTCTCGCGGGATTGACGGGCGCTCTCGAGAGCCGGCATCTCCTAGGAGCTAGTCTGGTC
CTCGGCTAGGCGGCTTGGGGTCCGCGCGTAACGGGGAGCCAGCCTGACGCGGGCGGACCCCGCTGTGATCCTG
GCAACGATGATGACTTGATGTTGGCACTGCGGCTTCCAGGAGGAGTGGAACCTGCAGGAGGCGGAGCGCGAT
CATGCCAGGAGTCCCTGTGCTAGTGGACGCGCTCGTGGGAGTTGGTGGACCCACACCGGACTTGCAGGCACTG
TTTGTTCAGTTTAAAGCAATCTTCTGGGGCCAGCTGGAGGCGCTCGAGGTGAAGTGGAGCGTGCGAATGACC
CTGTGTGCTGGGATATGCAGCTATGAAGGGAAGGTTGGAATGTGTTCCATCCGTCTCAGCGAACCCCTTTTGAAG
TTGAGGCCAAGAAAGGATCTTGTAGAGACCTCCTGCATGAAATGATACATGCCTATTTATTTTGTCACTAATAAC
GACAAAGACCGAGAAGGGCATGGTCCAGAATTTGTAAACATATGCATCGCATCAACAGCCTGACTGGAGCCAAAT
ATAACGGTATACCATACTTTTACGATGAGGTGGATGAGTATCGGCGACACTGGTGGCGCTGCAATGGGCGTGC
CAGCACAGGCCACCGTATTACGGCTATGTCAAACGAGCTACTAACAGGGAACCCCTCTGCTCATGACTATTGGTGG
GCTGAGCACCAGAAAACCTGTGGAGGCACTTACATAAAAAATCAAGGAACAGAGAATTACTCAAAAAAAGGCAAA
GGAAAGGCAAACTAGGAAAGGAACAGTATTGGCCGAGAGAATAAAGGTACCTTCGTGATATTCTTCTGATT
TTTTTGTGACCATAGCTATGATGTAAGACAATACTGTCTTCCAGAGAACTGGTATTAAAGATAAACTAAGGATC
GTTTCTGGTGTAGAACTCTTCAAGTGTAGACTTAAGGAAAAAATCCCACTGTCCATGAAATGATGGTAGGAAAC
AGACTTTGCTCTGTACAGAAGTAAGTAAAAGTAGGAATAGTTTCCATGGATATTTTTATTTTATTAACCTTTTTT
CAGTTTCTTTTATTCAAAGAAACAAATTTCAATCTCTGATAATATTGAGGTAAAGTTCCTTTCCCTATCTTGA
CTCACTGAGTTATTAGGAAACAGAAGGCAAAAGATTGTCAAATAAAAACAATAATTCAAGTAACAATGCCCGG
AATATACGTCCTAACTACACCCCTTCTATCAGCTGGATTCTATCCAAGTACTCTATTGATGTATGTATGTTCA
TTCAAAGAAATGGGAAAAGGATATGACATATATTTGCCAGTACTTCATCTTCAAGATTTACCCTTTTCTGTGAAG
TTCAGAGTTACTGAAGATGCTTCTTCCCTTGGGAAGTTGTTGACCCAAGAACATAGGTTATATTTCCCAAATCTT
TAATTATTGAGTGAAAGAGCTATAGATGAATTGATATGGAAAGACCGTATCTTCATTTTCGTGAGTAGAAGGAAA
GATAAGAATGAGGCAGCAGATTTTCCCTCCTGGAATTACACATAAAGGACACTAAGCAATTTTCAAGGTAATGT
TGCTTGTGTTGGTCTTTGGCATGATAAGATTCTTTATTTAAATATGAGAGAATTTTTTTTTATCCTTTATAT
CTCTCAATATCAGAACTCCTGAATTCTGAAGATTGCCCTCCTCCCATTAATAGGATTGTATGGATGTAAGATGGA
ATAAAATACTAGTTCTTCATTTTGAAGAACTGTACATTAGTTTAAATGTTTGTACTGTATTTCTTTTGAGTTGA
GGCACTTACATAACAATCTTCTTTGCTTTTTTGGCAGATAAACCCAAACAGAGGTGAGGCCAGCTAGTAATCCCT
TTTAGTGGGAAAGGATATGTTCTAGGAGAAACAGCAATTTACCTTACCTGGGAACTGATCACTTACATGCC
ATTAATAAAACCAAGATCTTTAAATCAAAACCATTCAGCAATGCTGTAAGACCTAATTCTAAAATCAAGGTG
AAATTTGAACAGAATGGTTCAAGTAAAAATTTCTACTGGTCTCCCTGCTGTTAGTAACAGTCACCAAAATGTT
CTAAGCAACTACTTTCTAGAGTATCATTTGCCAACCAAAAGGCTTTAGAGGTGTGAATGGATCTCCAAGGATA
AGTGTAACAGTTGGCAACATCCCTAAAACTCAGTCTCTTCTAGTTCTCAGAGAAGGGTTTCATCTTCTAAGATA
TCCCTAAGAAATTTCTTCAAAAGTAACGGAATCAGCATCTGTGATGCCATCCAGGATGTGAGTGGGTCTGAAGAT
ACATTTCCCAAATAAACGACCTAGGCTAGAAGATAAAAAAAA

78/550

FIGURE 78

MDDDLMLALRLQEEWNLQEAERDHAQESLSLVDASWELVDPTPDLQALFVQFNDQFFWGQLEA
VEVKWSVRMTLCAGICSYEGKGGMCSIRLSEPLLKLRPRKDLVETLLHEMIHAYLFVTNNDKD
REGHGPEFCKHMRINSLTGANITVYHTFHDEVDEYRRHWWRCNGPCQHRPPYYGYVKRATNR
EPSAHDYWWAEHQKTCGGTYIKIKEPENYSKKGKGKAKLGKEPVLAAENKGTFVYILLIFM

Important features of the protein:**Signal peptide:**

amino acids 1-41

N-glycosylation sites.

amino acids 148-151, 217-220

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 184-187

Casein kinase II phosphorylation sites.

amino acids 30-33, 121-124, 154-157, 187-190, 192-195

Tyrosine kinase phosphorylation site.

amino acids 211-218

N-myristoylation sites.

amino acids 59-64, 85-90, 146-151

Neutral zinc metallopeptidases, zinc-binding region signature.

amino acids 108-117

79/550

FIGURE 79

CGGACGCGTGGGTGGCAACCAGGAGAAGCCAAACTTGGTCCCCCGGCTCGCGGAGTGCCTGCG
AGCGGTGCTC**ATG**GCGCTCTATGAGGTCTTCTCTACCCGGTCGAGCGCAGTTACCGCGCGGG
GCTCTGCTCCAAAGCCGCGCTGTTCTGCTGCTGGCCGCTGCGCTCACGTACATCCCGCCGCT
GCTGGTGGCCTTCCGGAGCCACGGGTTTTGGCTGAAGCGGAGCAGCTACGAGGAGCAGCCGAC
CGTGCCTTCCAACACCAGGTGCTGCTCGTGGCCCTGCTCGGACCCGAAAGCGACGGGTTCTCT
CGCCTGGAGCACGTTCCCCGCTTCAACCGGTGCAAGGGGATCGCCTGCGCGTCCCGCTCGT
TTCGACTAGAGAAGAAGACAGGAACCAGGATGGGAAGACGGACATGTTACATTTTAAGCTGGA
GCTTCCCTGTCAGTCCACGGAGCACGTTCTCGGTGTGCAGCTCATCCTGACTTTCTCCTATCG
ATTACACAGGATGGCGACCCTCGTGATGCAGAGCATGGCGTTTCTCCAGTCTCCTTTCTGT
CCCGGGATCCCAGTTATACGTGAACGGAGACCTGAGGCTGCAGCAGAAGCAGCCGCTGAGCTG
TGGTGGCCTAGATGCCCCGATAACAATATCCGTGATCAACGGGACCAGCCCCCTTGGCCTATGA
CTACGACCTCACCCATATTGTTGCTGCCTACCAGGAGAGGAACGTTACCACCGTCCCTGAATGA
TCCCAACCCCATCTGGCTGGTGGGCAGGGCCGAGATGCTCCATTTGTGATTAATGCTATCAT
CCGATACCCTGTGGAAGTCATTTCTTATCAGCCAGGATTCTGGGAGATGGTAAAGTTCGCTG
GGTACAGTATGTCAGCATCCTGCTTATCTTCTCTGGGTGTTTGAAAGAATCAAGATCTTCGT
GTTTCAGAATCAGGTGGTGACCACCATTCCTGTGACAGTGACGCCCCGGGGAGACTTGTGTAA
GGAGCACTTATCT**TAG**AAAGGCCATTTCTGAAGACTCAGCAGGACCGTGGCTGCCTCATTGTC
ATCTTCTGGGAACATCTTAGGACCTTTTGAAAGAGCCAGCGGACACCTGCGGGCTTGTGTGC
TTTTCCCTCAGAGACAACGGTTCTTTCCGGTTTTGCTCTACACAGTTCCGTATCTTCAGAGCT
CCTGCAGAATTGTCAGGGACTAGTTTGTGGAAAGGTCTGAGAGTTCCTGGAGGCTATAATTAG
CTTTTGGGTTTTCTTCTTTGCTTAGCGTTGAATTTAGGAGAAAATTGCAGTCAGTTCAG
ACATCTTGGAAGAGTCCCATCTCTGGTCAAGCAGAGACTTTTCTCTGTTGAACTGAGGAAC
AACTGTGCATTTCTTCTTCTGTTGTGAGCCACTTTACTCTTTTCAGGGCTCTCTGTGAC
AAACATGCCAATCACTAGCACTTTGCACCCCTGGGCTTCTCCATTTCCCATTCACAGCTTTGA
TTTCCAGAGCTGAGGCCTTTAACTGGAGACCTGGAGGGGCAGGGCCCAAGGGCAAGGGCCGCA
TTAGCACAGGCAATCAGGGAGGGCCGCTGAAGGACACTTGGACCGTCCACCTGCCCCAGCCCA
ACAGTCAGTCATCTGTCATCAGCTCAGCTGAGCAGCCCTGGATCTTTGCCGTACTGTGACTGG
GCTCTTTGCCCTATTTTTCCCTCTGTCTGTGCCCTGGATGGCAGGCTGAAGTCAGAGGGGCT
GTTTCATTCTCAGCCCCCTCAGCAGCACTGGGGGAAGAAAGCATTGTCAACAGGTTCTTTC
TGGCCCTCACCAACAGCCTGGGCACTTGGCCCTCCTCCTCCTTGACAGCCCTCCCCCTTCT
GCAAAGGACAGGGGCGACAGGGGTTGGTGTGGGATTGGCTCCCGCTGCCTGACAACCACAAG
TTTATTTGGAAGGCTAGCGGGAAGCCAGCGGTGGCGTTTCCCTTGACTAAGGAACAGGGTG
CCCATCAGAGTGGGGCGGGCAGCTTTGGGAAGGACACAAGAAGCAGTAAGAGTGTAAGAGGA
TGCTGGCCTGGGCAGGCCAGTCCAGCCTGGCCACTAGCAGAATACCAAGCAGTCCAGTGGATT
ACCTCGTGGCTAAGCAAGTGTCTGCAGGAGCAGAGATGGCTGGAAGGGGCCTCTGCACACGG
AAGATGGCTTGTTCAGCCCATTACCTCCTGAGGATGTGGGCAGTCTCCTCCAAGAACACATG
GAGCTGCTTCTGATCCCAAGCAGGTCAATGCCACTGGAAGGACATGGCCCCGGTGATCCATG
CTTCATGCCCACCCAGAAACACACCCCTCAGTGTGTGCTCAGTTTACTTTGGAGATCAGTTG
TCGTTTTTAGTGCTCCTTTAGGCTTACTAAAACAGTTTTGGAAACAAAGCTATTTTGAAGTAT
TCAAGCAGAGGAATCCCTAACACTGACCCCTTGTCTTTTTTTAATATTCAGGCTGTTTTAT
ATGCCTAAATTTTTTCTTAAGATCTAAACGAAAAATAGTTTCTTGTTTAAATTCACATAAGG
CAATGAGATATGGAAAGATGACAAGATACGTATAAACATTGGTTTGCATCTTATTAAATTATT
CTAATGCAAATCTTGTATAAAGAACCCATGATGTTTTGTAACCTTTCTAATTAATGTTCAAA
ATGAG

80/550

FIGURE 80

MALYEVFSHPVERSYRAGLCSKAALFLLLAALTYIPPLLVAFRSHGFWLKRSSYEEOPTVRF
QHQVLLVALLGPESDGFLAWSTFFAFNRLQGDRLRVPLVSTREEDRNQDGKTDMLHFKLELPL
QSTEHLVGVQLILTFSYRLHRMATLVMQSMFLQSSFPVPGSQLYVNGDLRLQQKQPLSCGGL
DARYNISVINGTSPFAYDYDLTHIVAAYQERNVTTVLNDPNPIWLVGRAADAPFVINAIIRYP
VEVISYQPGFWEMVKFAWVQYVSILLIFLWVFERIKIFVFQNVVTTIPVTVTPRGDLCKEHL

Important features of the protein:**Signal peptide:**

amino acids 1-34

Transmembrane domain:

amino acids 268-284

N-glycosylation sites.

amino acids 194-198, 199-203, 221-225

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 51-55

Tyrosine kinase phosphorylation site.

amino acids 250-259

N-myristoylation site.

amino acids 187-193

Cell attachment sequence.

amino acids 307-310

81/550

FIGURE 81

GCCGGGAGCTTCCCTGATGGTGCCGCCGCCTCCGAGCCGGGGAGGAGCTGCCAGGGGCCAGCTGGGCAGGAGCCT
GGGTCCGCTGCTGCTGCTCCTGGCGTTGGGACACACGTGGACCTACAGAGAGGAGCCGGAGGACGGCGACAGAGA
AATCTGCTCAGAGAGCAAAATCGGACGACTAAATACCCGTGCTGAAGTCTTCAGGCGAGCTCACCACATGCTA
CAGGAAAAAGTGCTGCAAGGATATAAATTTGTTCTTGGACAATGCATCCAGAAGATTACGACGTTTGTGCCGA
GGCTCCCTGTGAACAGCAGTGCACGGACAACCTTTGGCCGAGTGTGTGTACTTGTATCGGGATACCGATATGA
CCGGGAGAGACACCGGAAGCGGGAGAAGCCATACTGTCTGGATATTGATGAGTGTGCCAGCAGCAATGGGACGCT
GTGTGCCCCACATCTGCATCAATACCTTGGGCAGCTACCGCTGCGAGTGCCGGGAAGGTACATCCGGGAAGATGA
TGGGAAGACATGTACCAGGGGAGACAAATATCCCAATGACACTGGCCATGAGAAGTCTGAGAACATGGTGAAGC
CGGAACCTTGCTGTGCCACATGCAAGGAGTTCTACCAGATGAAGCAGACCGTGCTGCAGCTGAAGCAAAAGATTGC
TCTGCTCCCCAACATGCAGCTGACCTGGGCAAGTATATCACTGGTGACAAGGTGCTGGCCTCAAACACCTACCT
TCCAGGACCTCCTGGCCTGCCTGGGGGCCAGGGCCCTCCCGGCTCACCAGGACCAAAGGGAAGCCCAGGCTTCCC
CGGTATGCCAGGCCCTCCTGGGCAGCCCGGCCACGGGGCTCAATGGGACCCATGGGACCATCTCCTGATCTGT
CCACATTAAAGCAAGGCCGAGGGGCCCTGTGGGTCCACAGGGGCACAGGAAGAGATGGTTCTAAGGGGGAGAG
AGGAGCGCCTGGGCCAGAGGGTCTCCAGGACCCCTGGTTCTTTGACTTCTGCTACTTATGCTGGCTGACAT
CCGCAATGACATCACTGAGCTGCAGGAAAAGGTGTTGGGCACCGGACTCACTCTTCAGCAGAGGAGTTCCCTTT
ACCTCAGGAATTTCCAGCTACCCAGAAGCCATGGACCTGGGCTCTGGAGATGACCATCCAAGAAGAAGTGAAGC
AAGAGACTTGAGAGCCCCAGAGACTTCTACCCATAGCACATCCCAACACCGTCACGCCAAAGGAAGAGAAAGAT
CAACTCACCTGCAGTTAAACCATCTAAAGAGAAGAAAGACCACTGGAGACCTAGAAAACATACATTTTCTCTTC
GCTTCTCCTGACGTCTCTCCACTCCTTCTTCCAAATACGATGCTATTTTCAGAGTCCCTCCTAGGCCTGCAG
ACATGAGGGAGTGAATGATTGATTTACCTGCTTCTCACTAAGAGTCCATTGGGGTGGTTGCATTGTAACCTTTTC
TTTTACATCCTATTTTCCAGGAACCTTGGATTTAAGTACTCTCACAGTGTCTTAAATCATAAATTTCTGAAGTT
AAATTTGGCAGAGTATCAAAAGGGGGAAAATGACAAAGTGAGCTCTAAGAAAATGTGAGGCTACTTCTAAGATT
GTGTTCACAATAGACCATAACTCCTCTAGTATCAAAATTTGGGGCTCTTCAGTTAAAAAGGGTGGGGAGGACAAA
CGGTGCGATGTGCTTTGGTGAGAAATTTTTCCTTGTGCTTCTAGTAGACTTTAAATATTGTATCCCTTTGTCAA
ACCTTGTTTTCCCAAAATCAATTAAGAGAGGAGAGAATTGAATGGCGTTTAGAGAAGATAGAAAAGAATCACAGT
CATATATTTACTGTTATATAGATTGCCACATTCTAAAATTCAAATACGGTGCTTAAGGTTTCATGCCATGCTTAT
CTGTAAGTATCCTATTTAGGGAAGAAGATTAACTCTCTTTCAAAAAACAAAGTGAATGCCTGGATTACAT
TAAACAATGGGCTCTCGTTTGCTATAATATTTTAAAGCTGTTTAAATCAACAGTGGAGTCTGCTCTATAAATATA
GATTATTTGTTCAATAAACTGGCTGAGCTTAGAGAGAGGTGCAGAATCCTGGTCTGAGCAGGTGCCAGAAGG
TACCATTAGGTGCCATGATCCAGGCTGAACCAATATACAGTGGGGCTGAAGTCTGCAAGGAGGTTGCTGGCTTGG
GCTGACCTCACTAATGCCATCAGCAGCGGTAGGTAAATTTTTCTCCTTGGGTATTACAAGTTTTGTCTGGAGC
CAACCAAGCTTGCCACCAACATATTGAGAGTAATACACTATTGAAAGTTATCTTGGATGGGGAGAAAAAAATA
GTGGTTTTCTTGTGTTTGCAAAAACCTTCTTCTATTCTCATTTTTTCTTAATTTTCTTTAATTTAGTCCAAGTTC
CAGTTCTTTTAGGCCTTCTCTTTGATTTATTTTCCCCTGCATGTGAGAAGCAGTTTCAAAAAAGGTCTATATCTC
CACCTCCTAGTGAGTTAGAGTGTCTTCTCAGAGCACCTTGGGTGGCAAAGGGAAGCATGTTCTGCCAAGGTTT
GCTGTGGATTCAAGACACCAGGAGCAAGAGACCAGAAGGATGATCTGCTCCTTTGTAACGTTGTTGAGGGCCCT
CTTGTTCCAATGAGCAGCTTATAGGTTACTCACAGTCCACTTCTCACTGGACACACAAAGTGGCTCTTTATCT
ACCTTTGCGGGAGATTTTCACTCTCCTGCAATGATCGTTCTCACACTCATATTAGCTCATGTTGGAATTTCCCA
TCCTGCCATGTCCTTTCCCATTTCTTTTGGCTTTTTTGCTCCACCTTTTAGCCACATCATTTAACTCCACTA
CTGTGAAAGCTTGCTTAAAGAAAATCCCTCTTGGCCGGGTGTGGTAGCCACGCCTCTAATCCCAGCACTTTGGG
AGGCTGAGGCGGGGAGATCAAAAGTCCAGGAGATCGAGACCAGCCTGACCAACATGGTGAACCCCTGTCTCTACT
AAAAATACAAAATTAGCTGGGCGTGTGGCACACACCTGTAAATCCAGCTACTCAGGAGGCTGAGGCAGGAGAA
TTACTTTAACCTGCGGGGGAGCCTAGATTGCGCTACTGCACTCCAGCCTAGGCAACAGAGGGAGACTCTGTCTC
ATTAAAAA

82/550

FIGURE 82

MVPPPPSRGGAARGQLGRSLGPLLLLLALGHTWTYREEPEDGDREICSESKIATTKYPCLKSS
GELTTCYRKCKCKGYKFVLGQCIPEDYDVCAEAPCEQQCTDNFGRVLCTCYPGYRDRERHRK
REKPYCLDIDECASSNGTLCAHICINTLGSYRCECREGYIREDDGKTCTRGDKYPNDTGHEKS
ENMVKAGTCCATCKEFYQMKQTVLQLKQKIALLPNNAADLGKYITGDKVLASNTYLPGPPGLP
GGQGPPGSPGPKGSPGFPGMPGPPGQPGPRGSMGPMGSPDL SHIKQGRRGVPGPPGAPGRDG
SKGERGAPGPRGSPGPPGSFDFLLMLADIRNDITELQEKVFGHRTHSSAEFFPLPQEFFPSYP
EAMDLGSGDDHPRRTETRDLRAPRDFYP

Important features of the protein:**Signal peptide:**

amino acids 1-34

N-glycosylation sites.

amino acids 142-148, 182-188

Tyrosine kinase phosphorylation site.

amino acids 125-132

N-myristoylation sites.

amino acids 10-16, 143-149, 155-161, 196-202, 250-256

Amidation site.

amino acids 299-303

Aspartic acid and asparagine hydroxylation site.

amino acids 150-162

Cell attachment sequence.

amino acids 176-179

Clq domain proteins.

amino acids 247-280

Calcium-binding EGF-like domain proteins pattern proteins.

amino acids 144-165

83/550

FIGURE 83

ATCTGAGTGAGCTAACTGACACAATGAAACTGTCAGGCATGTTTCTGCTCCTCTCTCTGGCTC
TTTTCTGCTTTTTTAACAGGTGTCTTCAGTCAGGGAGGACAGGTTGACTGTGGTGAGTTCCAGG
ACCCCAAGGTCTACTGCACTCGGGAATCTAACCACACTGTGGCTCTGATGGCCAGACATATG
GCAATAAATGTGCCTTCTGTAAGGCCATAGTGAAAAGTGGTGGAAAGATTAGCCTAAAGCATC
CTGGAAAATGCTGAGTTAAAGCCAATGTTTCTTGGTGACTTGCCAGCTTTTGCAGCCTTCTTT
TCTCACTTCTGCTTATACTTTTGCTGGTGGATTCCTTTAATTCATAAAGACATACCTACTCTG
CCTGGGTCTTGAGGAGTTCAATGTATGTCTATTTCTCTTGATTCACTTGTCAATAAAGTACATTC
TGCAAAAGCAAAAA

84/550

FIGURE 84

MKLSGMFLLLSLALFCFLTGVFSQGGQVDCGEFQDPKVYCTRESNPFCGSDGQTYGNKCAFCK
AIVKSGGKISLKHGPGKC

Important features of the protein:

Signal peptide:

amino acids 1-23

N-myristoylation sites.

amino acids 26-32, 52-58, 56-62, 69-75

Kazal serine protease inhibitors family signature.

amino acids 40-63

85/550

FIGURE 85

GGAGCAGACACACAGACCCGGGCCGGAGGCCCTCTTCTAGCCCTGCGGGAACCGGACAGTTC
CCCAACTGGGGACTCTGGAACCACAGCTCCTAAATCATCAAATTTCTCAAGCTTTTTTTTTCCC
TCTCTTCGTCCCAGCCATCCCAGTCTTCTTCTTCTTTTTTTTTTTTTTAAGTTATTGTTTTT
TCGCTCCTGTCAATTATGAAAGTGGTCACGCCATTCAATATTAAGACTTGGAGGGAATTGGGGA
AAGAAAAGAAAGAATCTAAAAGAAGAGAAGCGACCGGTGCTTTTAAGGGTGTCTAATTTTCAA
AAGAGACGTCTGGGAGTATTTTGCTCTGGGCGTTTGGAGCAACTTCGCGGACAGCGGAGCTCG
CCCAGCATGGATGTTCCAGGTTACAGGCGCCTTCTTCTGAGAACGACCCTGGCCTTGAACG
TCAGAGCCGGGGACGAAGGCCCCGGAGGCTGCTGCGAGCTCCGCGCGTTCTTCGCGCCCTT
CCGCGCCGCTCGCGCCGGCGCCGGCCTCCACCCCGCGCGCCGCTCCACAGTCCCGATGC
AGGCGCCCGGGCCGGGGGCCACTCGGGCTGCGGCTGATGATGCCCGGGCGCCGGGGGGCGCTGC
GCGAGCCTGGCGGCTGCGGATCCTGCCTGGGGGTGGCGCTGGCCCTGCTGTTGCTGCTACTGC
CCGCCTGCTGCCCCGTGCGGGCGCAGAACGACACGGAGCCCATCGTGCTGGAGGGCAAGTCCG
TGGTGGTGTGCGACTCCAGCCCGTCGGCGGACGGCGCCGTCACCTCCTCCCTAGGCATCTCCG
TGCGCTCCGGCAGCGCAAGGTGGCCTTCTCCGCCACGCGGAGCACCAACCACAGAGCCGTCCG
AGATGAGCAACCGCATGACCATCTATTTTCGACCAGGTATTAGTAAATATTGGCAACCACT
TTGATCTTGCTTCCAGTATATTTGTAGCACCGAGAAAAGGGATTTATAGCTTCAGCTTCCACG
TGGTCAAAGTGTATAACAGACAAACCATCCAGGTCAGTTTAATGCAGAATGGCTACCCAGTGA
TCTCGGCCTTTGCAGGAGACCAGGATGTACCAGAGAAGCTGCTAGCAATGGCGTGCTGCTGC
TCATGGAAGGGGAAGACAAAGTGCATCTCAAACCTTGAGAGAGGCAACCTCATGGGGGGCTGGA
AATACTCCACATTCTCGGGCTTCTTGGTGTTCCTCTATAACACAGAGCCCCCTAGATGGTG
GGGGAATGGCAAACCTGGACCCAGGACTCCGCCCTTTAAAACACCCTGAACCTTACTGGAATTGG
ACACCTTGTTCACACCTCCGTCAGACTGTTGCAGTAGAAGAATGATTTCCCTTGAACCTCC
AGTACTTTTGTTCCTTGTTCCTTGTTCCTTGTTCCTTGTTCCTTGTTCCTTGTTCCTTGTTCCT
TTTAGATGACAAGGTCTTAAGGAGAAATGAAATTATCGATTGAGCAATTTGTACCTGTGATT
GTAAAGTCAATATCGGATTTTATTGTTGGGACCATGGACCTCTTTTGTTCCTTGTTCCTTGTTCCT
TCGTCCCAACGGAAGGAGAGCTCCTGACTCCAGGATGGGCTGCAGGTTGCAGTCAGGGCTTGA
AGTAGGAGCCAGCAAAGAACCACCTGCTGGACAGTCCTTGACATGTGTTCTGTGTGTGTCTG
TATAGCCTTAAGAAAAGAATGGCTTCACTTTCACTTCTGTATTCTTCCCCCACCATGTGGCT
GGGAGGACTTGGGAGGGGGATGGGGACATTGGGAACCTGTCAAGAAGTGCTTTATCCAGAGAA
GCAAATTTTGCACGATTGGACTGCAATTTTGTTCCTTGTTCCTTGTTCCTTGTTCCTTGTTCCT
CTTTACTTTTCTTTCACACTCAGCTCTCCCTCCTCAACCCCACTTTTATTTTCTTGTTCCTTGTTCCT
GTTGAGGAGAGAAAATATAGAATTCCTGGATAAGACCAAACAAAACAAACATTAATAACCT
GTATGTTTTGTTTTAGACGAGACCAACTAAACAAAAGTATCTGTTTATCAAAGTAAAAGTA
ACACAATGGACAATTCTGCTTATCTCTCAAAGAGATTCTAAGATGCACCTTTAGAACTATTA
ATAGCAACCTGCATTTTTTTTTTAATTTATACTTCAGAATCCTTTAAGAACCTGGTGTTCCTGA
GTGGTCTGAATCATATAAGTTGGTAATGGAAGCTGTAATGACCAAGTCCCCTAAACATACTA
TGTCTTTGCCACGTGTGCTGTGACTTCTCTGTGGGTGATTTAATTTATTTGGATCCACCTCTG
AGTGAGCGCACAGTGATCAGGTGCTTCAAAGCCAACAGACCAGCTCCTCTTCCCTCCGGATCCT
CTTTTGATCTGCCAGGAAAGGGATGCATTGACACTCTCCTGCATGCACCTGGCGAGAAGCCA
CCTGAAAGTCACTGTGGTTAAAGATATTGGTGGAGGTACCCAGGAGCACTGTTACAAATCCT
TCTTGTTCCTTGTTCCTTGTTCCTTGTTCCTTGTTCCTTGTTCCTTGTTCCTTGTTCCTTGTTCCT
CATATGCCAAGGAAATGTCACTAATCCCAAAGCAATCAAAAAGGAGACCTCAAACAGATGTT
AATTTGTTCTTTGTGTAACAATGTAACCAAAATATTGATGATAAAGTCATAATTTAAGATTC
AGAATAAATGGGTTTGATGTCTGGCAAAAAAAAAAAAAA

86/550

FIGURE 86

MQAPGRGPLGLRLMMPGRRGALREPGGCGSCLGVALALLLLLLLPACCPVRAQNDEPIVLEGK
CLVVCDSPPSADGAVTSSLGISVRSGSAKVAFSATRSTNHEPSEMSNRTMTIYFDQVLVNIGN
HFDLASSIFVAPRKGIIYSFSFHVVVKVYNRQTIQVSLMQNGYPVISAFAGDQDVTREAASNGVL
LLMEREDKVHLKLERGNLMGGWKYSTFSGFLVFPL

Important features of the protein:**Signal peptide:**

amino acids 1-48

N-glycosylation sites.

amino acids 53-57, 110-114

N-myristoylation sites.

amino acids 26-32, 27-33, 29-35, 33-39, 76-82, 205-211

Amidation site.

amino acids 16-20

Clq domain signature.

amino acids 117-148

Clq domain proteins.

amino acids 115-149

87/550

FIGURE 87

AGGGCCCGCGGGTGGAGAGAGCGACGCCCCGAGGGGATGGCGGCAGCGTCCCGGAGCGCCTCTG
GCTGGGCGCTACTGCTGCTGGTGGCACTTTGGCAGCAGCGCGCGGCGGCTCCGGCGTCTTCC
AGCTGCAGCTGCAGGAGTTCATCAACGAGCGCGGCGTACTGGCCAGTGGGCGGCCTTGCGAGC
CCGGCTGCCGGACTTTCTTCCGCGTCTGCCTTAAGCACTTCCAGGCGGTCTGCTCGCCCCGGAC
CCTGCACCTTCGGGACCGTCTCCACGCCGTATTGGGCACCAACTCCTTCGCTGTCCGGGACG
ACAGTAGCGGCGGGGGGCGCAACCCTCTCCAAGTGCCTTCAATTTACCTGGCCGGGTACCT
TCTCGCTCATCATCGAAGCTTGGCACGCGCCAGGAGACGACCTGCGGCCAGAGGCCTTGCCAC
CAGATGCACTCATCAGCAAGATCGCCATCCAGGGCTCCCTAGCTGTGGGTGAGAACTGGTTAT
TGGATGAGCAAACCAGCACCCCTCACAAGGCTGCGCTACTCTTACCGGGTCATCTGCAGTGACA
ACTACTATGGAGACAACTGCTCCCGCCTGTGCAAGAAGCGCAATGACCACTTCGGCCACTATG
TGTGCCAGCCAGATGGCAACTTGTCTGCTGCCCCGTTGGACTGGGGAATATTGCCAACAGC
CTATCTGTCTTTCGGGCTGTCATGAACAGAATGGCTACTGCAGCAAGCCAGCAGAGTGCCTCT
GCCGCCAGGCTGGCAGGGCCGGCTGTGTAACGAATGCATCCCCACAATGGCTGTGCCACG
GCACCTGCAGCACTCCCTGGCAATGTACTTGTGATGAGGGCTGGGGAGGCCTGTTTTGTGACC
AAGATCTCAACTACTGCACCCACCACTCCCCATGCAAGAATGGGGCAACGTGCTCCAACAGT
GGCAGCGAAGCTACACCTGCACCTGTGCGCCAGGCTACACTGGTGTGGACTGTGAGCTGGAGC
TCAGCGAGTGTGACAGCAACCCCTGTGCAATGGAGGCAGCTGTAAGGACCAGGAGGATGGCT
ACCACTGCCTGTGCTCCTCCGGGCTACTATGGCCTGCACTGTGAACACAGCACCTTGAGCTGCG
CCGACTCCCCCTGCTTCAATGGGGGCTCCTGCCGGGAGCGCAACCAGGGGGCCAACTATGCTT
GTGAATGTCCCCCAACTTCACCGGCTCCAAGTGCAGAGAAGAAAGTGGACAGGTGCACCAGCA
ACCCCTGTGCCAACGGGGGACAGTGCCTGAACCGAGGTCCAAGCCGCATGTGCCGCTGCCGTC
CTGGATTACGGGCACCTACTGTGAAGTCCACGTCAGCGACTGTGCCCGTAACCCTTGCGCCC
ACGGTGGCACTTGCCATGACCTGGAGAATGGGCTCATGTGCACCTGCCCTGCCGGCTTCTCTG
GCCGACGCTGTGAGGTGCGGACATCCATCGATGCCTGTGCCTCGAGTCCCTGCTTCAACAGGG
CCACCTGCTACACCGACCTCTCCACAGACACCTTTGTGTGCAACTGCCCTTATGGCTTTGTGG
GCAGCCGCTGCGAGTTCCCCGTGGGCTTGCCGCCAGCTTCCCCTGGGTGGCCGTCTCGCTGG
GTGTGGGGCTGGCAGTGCTGCTGGTACTGCTGGGCATGGTGGCAGTGGCTGTGCGGCAGCTGC
GGCTTCGACGGCCGGACGACGGCAGCAGGGAAGCCATGAACAACCTTGTCCGACTTCCAGAAGG
ACAACCTGATTCTGCGCCCGAGCTTAAAAACACAAACCAGAAGAAGGAGCTGGAAGTGGACT
GTGGCCTGGACAAGTCCAAGTGTGGCAAACAGCAAAACCACACATTGGACTATAATCTGGCCC
CAGGGCCCCCTGGGGCGGGGGACCATGCCAGGAAAGTTCCCCACAGTGACAAGAGCTTAGGAG
AGAAGGCGCCACTGCGGTTACACAGTGAAAAGCCAGAGTGTCCGATATCAGCGATATGCTCCC
CCAGGGACTCCATGTACCAGTCTGTGTGTTGATATCAGAGGAGAGGAATGAATGTGTGCTTGC
CCACGGAGGTATAAGGCAGGAGCCTACCTGGACATCCCTGCTCAGCCCCGCGGCTGGACCTTC
CTTCTGCATTGTTTACA

88/550

FIGURE 88

MAAASRSASGWALLLLVALWQQRAAGSGVFQLQLQEFINERGVLASGRPCEPGCRTEFFRVCLK
HFQAVVSPGPCTFGTVSTPVLGTNSFAVRDDSSGGGRNPLQLPFNFTWPGTFSLLIEAWHAPG
DDLRLPEALPPDALISKIAIQGSLAVGQNWLLDEQTSTLTRLRYSYRVICSDNYYGDNCSRLCK
KRNDHFGHYVCQPDGNLSCLPGWTGEYCQQPICLSGCHEQNGYCSKPAECLCRPGWQGRCLNE
CIPHNGCRHGTCSTPWQCTCDEGWGGLFCDQDLNYCTHHSPCKNGATCSNSGQRSYTCTCRPG
YTGVDCELELSECDNPCRNGGSKDQEDGYHCLCPPGYGLHCEHSTLSCADSPCFNGGSCR
ERNQGANYACECPPNFTGSNCEKKVDRCTSNPCANGGQCLNRGPSRMCRCRPGFTGTYCELHV
SDCARNPCAHHGTCHDLENGLMCTCPAGFSGRRCVVRTSIDACASSPCFNATCYTDLSTDTF
VCNCPYGFVGSRCFVPVGLPPSFPWVAVSLGVGLAVLLVLLGMVAVAVRQLRLRRPDDGSREA
MNNLSDFQKDNLI PAAQLKNTNQKKELEVDCGLDKSNCGKQQNHTLDYNLAPGPLGRGTMPGK
FPHSDKSLGEKAPLRLHSEKPECRISAICSPRDSMYQSVCLISEERNECVIATEV

Important features of the protein:**Signal peptide:**

amino acids 1-26

Transmembrane domain:

amino acids 530-552

N-glycosylation sites.

amino acids 108-112, 183-187, 205-209, 393-397, 570-574, 610-614

Glycosaminoglycan attachment site.

amino acids 96-100

Tyrosine kinase phosphorylation site.

amino acids 340-347

N-myristoylation sites.amino acids 42-48, 204-210, 258-264, 277-283, 297-303, 383-389,
415-421, 461-467, 522-528, 535-541, 563-569, 599-605, 625-631**Amidation site.**

amino acids 471-475

Aspartic acid and asparagine hydroxylation site.

amino acids 339-351

EGF-like domain cysteine pattern signature.amino acids 173-185, 206-218, 239-251, 270-282, 310-322, 348-360,
388-400, 426-438, 464-476, 506-518**Calcium-binding EGF-like:**amino acids 224-245, 255-276, 295-316, 333-354, 373-394, 411-432,
449-470

89/550

FIGURE 89

GTCTCCGCGTCACAGGAACCTTCAGCACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGAC
TTGACTCCCGCGCGCCCCAACCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGC
CGCCAGTCCCGGCCCTCTCCCGCCCCACACCCACCCTCCTGGCTCTTCCTGTTTTTACTCC
TCCTTTTCATTATAACAAAAGCTACAGCTCCAGGAGCCCAGCGCCGGGCTGTGACCCAAGCC
GAGCGTGGAAGAATGCGGGTTCCTCGGGACCGGCACTTGGATTCTGGTGTAGTGCTCCCGATT
CAAGCTTTCCCAAACCTGGAGGAAGCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCA
GAAAGACCTTTGAATGAACAGATTGCTGAAGCAGAAGAAGACAAGATTAAAAAACATATCCT
CCAGAAAACAAGCCAGGTCAGAGCAACTATTCTTTTGTGATAACTTGAACCTGCTAAAGGCA
ATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTATAAGAAGCTCCCCACTTGATAAT
AAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACGATCGATGATTATGACTCT
ACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCATCAACTAGACGGG
ACTCCTTTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTTATGAAGAAAATGAC
AGAGCCGTGTTTGACAAGATTGTTTCTAAACTACTTAATCTCGGCCTTATCACAGAAAGCCAA
GCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAAATTAATCTCAAAGGAAGCCAAC
AATTATGAGGAGGATCCCAATAAGCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCA
GAGAAAGTGAAGTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACA
GTATCTAACACATTAACCTTGACAAATGGCTTGGAAGGAGAACTAAAACCTACAGTGAAGAC
AACTTTGAGGAACTCCAATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTAGAA
AAAGAAGCAAAGAGAAAGAAACACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAG
ATGATGGTGAAATATGGAACAATATCTCCAGAAGAAGGTGTTTCCTACCTTGAAAACCTGGAT
GAAATGATTGCTCTTCAGACCAAAAACAAGCTAGAAAAAAATGCTACTGACAATATAAGCAAG
CTTTTCCCAGCACCATCAGAGAAGAGTCATGAAGAAACAGACAGTACCAAGGAAGAAGCAGCT
AAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCACAAAAGATGATAACTCCAACCCAGGA
GGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGCCATCAGAAAAAATATT
GAATGGTTGAAGAAACATGACAAAAGGGAAATAAAGAAGATTATGACCTTTCAAAGATGAGA
GACTTCATCAATAAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAAGAAGCC
GAGGCCATCAAGCGCATTTATAGCAGCCTGTAAATGGCAAAAGATCCAGGAGTCTTTCAAC
TGTTTCAGAAAACATAATATAGCTTAAACACTTCTAATTCTGTGATTAAAAATTTTTTGACCC
AAGGGTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGC
TTTCTTCCCGTAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAA
AAA

90/550

FIGURE 90

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENK
PGQSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSG
LDHKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTL
EDEVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIKEKVTMAAIQDGLAKGENDETVSNT
LTLTNGLERRTKTYSEDNFEELQYFPN FYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVK
YGTISPEEGVSYLENLDEMIALQTKNKLEKNATDNISKLF PAPSEKSHEETDSTKEEAAKMEK
EYGSLKDSTKDDNSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFIN
KQADAYVEKGILDKEEAEAIKRIYSSL

Important features:**N-glycosylation sites:**

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390**N-myristoylation sites:**

amino acids 143-148, 239-244

91/550

FIGURE 91

TGCATCAGTGCCCAGGCAAGCCCAGGAGTTGACATTTCTCTGCCAGCCATGGGCCTCACCCCT
GCTCTTGCTGCTGCTCCTGGGACTAGAAGGTCAGGGCATAGTTGGCAGCCTCCCTGAGGTGCT
GCAGGCACCCGTGGGAAGCTCCATTCTGGTGCAGTGCCACTACAGGCTCCAGGATGTCAAAGC
TCAGAAGGTGTGGTGCCGGTTCTTGCCGGAGGGGTGCCAGCCCCTGGTGTCTCAGCTGTGGA
TCGCAGAGCTCCAGCGGGCAGGCGTACGTTTCTCACAGACCTGGGTGGGGGCCTGCTGCAGGT
GGAAATGGTTACCCTGCAGGAAGAGGATGCTGGCGAGTATGGCTGCATGGTGGATGGGGCCAG
GGGGCCCCAGATTTTGCACAGAGTCTCTCTGAACATACTGCCCCCAGAGGAAGAAGAAGAGAC
CCATAAGATTGGCAGTCTGGCTGAGAACGCATTCTCAGACCCTGCAGGCAGTGCCAACCCCTT
GGAACCCAGCCAGGATGAGAAGAGCATCCCCTTGATCTGGGGTGCTGTGCTCCTGGTAGGTCT
GCTGGTGGCAGCGGTGGTGCTGTTTGCTGTGATGGCCAAGAGGAAACAAGAATCCCTCCTCAG
TGGTCCACCACGTCAGTTGACTCTGGACCGGCTGCTGAATTGCCTTTGGATGTACCACACATTA
GGCTTGACTCACCACCTTCATTTGACAATACCACCTACACCAGCCTACCTCTTGATTCCCCAT
CAGGAAAACCTTCACTCCCAGCTCCATCCTCATTGCCCCCTCTACCTCCTAAGGTCCTGGTCT
GCTCCAAGCCTGTGACATATGCCACAGTAATCTTCCCGGGAGGGAACAAGGGTGGAGGGACCT
CGTGTGGGCCAGCCCAGAATCCACCTAACAATCAGACTCCATCCAGCTAAGCTGCTCATCACA
CTTTAAACTCATGAGGACCATCCCTAGGGGTTCTGTGCATCCATCCAGCCAGCTCATGCCCTA
GGATCCTTAGGATATCTGAGCAACCAGGGACTTTAAGATCTAATCCAATGTCCTAACTTTACT
AGGGAAAGTGACGCTCAGACATGACTGAGATGTCTTGGGGAAGACCTCCCTGCACCCAACTCC
CCCACTGGTTCTTCTACCATTACACACTGGGCTAAATAAACCCTAATAATGATGTGCAAAAAA
AA

92/550

FIGURE 92

MGLTLLLLLLGLLEGQIVGSLPEVLQAPVGSSILVQCHYRLQDVKAQKVWCRFLPEGCQPLV
SSAVDRRAPAGRRTFLTDLGGGLLQVEMVTLQEEDAGEYGCMVDGARGPQILHRVSLNILPPE
EEEETHKIGSLAENAFSDPAGSANPLEPSQDEKSIPLIWGAVLLVGLLVAAVVLFAVMAKRKQ
ESLLSGPPRQ

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 161-181

N-myristoylation sites.

amino acids 17-23, 172-178

Amidation site.

amino acids 73-79

93/550

FIGURE 93

GGCGGCGTTGCCGGGCTCTCCGGAAGGAGACGTGGCGGCGGTTGGGCCGGTGATACCCGGGCG
CTTTATAGTCCCGCCGCCTCCTCCTCCACCTCCTCCTCCTCCTCCTCCTGCGGCAGAG
GAGGTTGTGGCGGTGGCTGGAGAAAGCGGCGGCGGAGG**ATG**GAGGAAGGAGGCGGCGGCGTAC
GGAGTCTGGTCCCGGGCGGGCCGGTGTACTGGTCCTCTGCGGCCCTCCTGGAGGCGTCCGGCG
GCGGCCGAGCCCTTCCTCAACTCAGCGATGACATCCCTTTCCGAGTCAACTGGCCCGGCACCG
AGTTCTCTCTGCCCACAACCTGGAGTTTTATATAAAGAAGATAATTATGTCATCATGACAACCTG
CACATAAAGAAAAATATAAATGCATACTTCCCCTTGTGACAAGTGGGGATGAGGAAGAAGAAA
AGGATTATAAAGGCCCTAATCCAAGAGAGCTTTTGGAGCCACTATTTAAACAAAGCAGTTGTT
CCTACAGAATTGAGTCTTATTGGACTTACGAAGTATGTCATGGAAAACACATTTCGGCAGTACC
ATGAAGAGAAAGAACTGGTCAGAAAATAAATATTCACGAGTACTACCTTGGGAATATGTTGG
CCAAGAACCTTCTATTTGAAAAAGAACGAGAAGCAGAAGAAAAGGAAAAATCAAATGAGATT
CCTACTAAAAATATCGAAGGTCAGATGACACCATACTATCCTGTGGGAATGGGAAATGGTACAC
CTTGTAGTTTGAACAGAACCCGGCCAGATCAAGTACTGTGATGTACATATGTCATCCTGAAT
CTAAGCATGAAATTCTTTCAGTAGCTGAAGTTACAACTTGTGAATATGAAGTTGTCATTTTGA
CACCCTCTTGTGCAGTCATCCTAAATATAGGTTTCAAGCATCTCCTGTGAATGACATATTTT
GTCAATCACTGCCAGGATCTCCATTTAAGCCCCCTCACCCTGAGGCAGCTGGAGCAGCAGGAAG
AAATACTAAGGGTGCCTTTTAGGAGAAATAAAGAGGGTGTGCGTTGGTGGAAATATGAATTCT
GCTATGGCAAACATGTACATCAATACCATTGAGGACAAGGATAGTGGGAAACCTCTGTGGTTG
TCGGGACATGGAACCAAGAAGAGCATATTGAATGGGCTAAGAAGAATACTGCTAGAGCTTATC
ATCTTGAAGACGATGGTACCCAGACAGTCAGGATGGTGTACATTTTTTATGGAAATGGAGATA
TTTGTGATATACTGACAAACCAAGACAGGTGACTGTAAACTAAAGTGCAAAGAATCAGATT
CACCTCATGCTGTTACTGTATATATGCTAGAGCCTCACTCCTGTCAATATATTCTTGGGGTTG
AATCTCCAGTGATCTGTAAATCTTAGATACAGCAGATGAAAATGGACTTCTTCTCTCCCA
ACT**TAA**AGGATATTAAAGTTAGGGGAAAGAAAAGATCATTGAAAGTCATGATAATTTCTGTCCC
ACTGTGTCTCATTATAGAGTTCTCAGCCATTGGACCTCTTCTAAAGGATGGTATAAAATGACT
CTCAACCACTTTGTGAATACATATGTGTATATAAGAGGTTATTGATAAACTTCTGAGGCAGAC
ATTTGTCTCGCTTTTTTTTCATTTTTGTTGTGTCTTATAAACTGACTGTTTTTCTTTGCTTGA
TACTGTGATTCCAAAATAAATCTCATCCAAGCAAGTTAGAGTCCAGCCTAATCAAATGTCATA
ATTGTTGTACCTATTGAAAGTTTTTAAATAATAGATTTATTATGTAAATTATAGTATATGTAA
GTAGCTAATGAAGTAAAGATCATGAAGAAAGAAATTGATAGGTGTAAATGAGAGACCATGTAA
AATATGTAAATTCTAGTACCTGAAATCCTTTCAACAGATTTTTTATATAGCAACTGCTCTCTGC
AAGTAGTTAACTAGAACTGGGCACATGGTAGAGGCTCACATGGGAGTTGTCCCTCACCCTTG
TTAATCTCAAGAACTCTTATTTATAATAGGTTGCTTCTCTCTCAGAACTTTTATCTATTACT
TTTTTCTTCTTATGAGTATGTTTACTCTCAGAGTATCTATCTGATGTAGACAGTTGGTGTATGC
TTCTGAGACTCAGAATGGTTTACTCTAACAAAACACTGTGCTGTCTATCCCTTGTAATTGCCT
ACTGTAATATGGATTTCACTTCTGAACAGTTTACAGCACAATATTTATTTTAAAGTGAATAAA
ATGTCCACAAGCAAAA

94/550

FIGURE 94

MEEGGGGVRS LVPGGPVLLVLCGLLEASGGGRALPQLSDDIPFRVNWPGTEFSLPTTGVLVLYKE
DNYVIMTTAHKEKYKCILPLVTSGDEEEEEKDYKGPNPRELLEPLFKQSSCSYRIESYWTYEV
HGKHIRQYHEEKETGQKINIHEYLG NMLAKNLLFEKEREAEKEKSNEIPTKNIEGQMTPYY
PVGMGNGTPCSLKQNRPRSSTVMYICHPESKHEILSVAEVTTC EYEVVILTPLLCSHPKYRFR
ASPVNDIFCQSLPGSPFKPLTLRQLEQQEEILRVPFRRNKEGVGWWKYEF CYGKHVHQYHEDK
DSGKTSVVVG TWNQEEHIEWAKNTARAYHLQDDGTQTVRMVSHFYGN GDICDITDKPRQVTV
KLKCKESDSPHAVTVYMLEPHSCQYILGVESPVICKILDTADENGLLSLPN

Important features of the protein:

Signal peptide:

amino acids 1-30

Glycosaminoglycan attachment site.

amino acids 28-32

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 337-341

N-myristoylation sites.

amino acids 6-12, 23-29, 29-35, 49-55, 141-147, 152-158, 192-198,
196-202

Gram-positive cocci surface proteins 'anchoring' hexapeptide..

amino acids 54-60

95/550

FIGURE 95

TTCCGTTTCTGGGAGGAGTGAGGGGCAACGGGTCGGAGAAAAAGGAAAAAGAAGGGCTCAGC
GCCTCCCCGCCGGCCGTGGACAGAGGGGCACAGTTTCGGCAGGCGGGTGAGGTCGCTGAGGG
CCCGCCGGAGATGTTTTCTTGTGCGAGCACGGTGCAACCCCAGGTTACAGTTCCTCTGAGTCA
TCTCATCAATGCCTTCCATACACCAAAAAACACTTCTGTTTCTCTCAGTGGAGTGTGAGTTTC
TCAAAACCAGCATCGAGATGTAGTTCCTGAGCATGAGGCTCCCAGCAGTGAGCCTTCACTTAA
CTTAAGGGACCTTGGATTATCTGAACTAAAAATTGGACAGATTGATCAGCTGGTAGAAAATCT
ACTTCCTGGATTTTGTAAAGGCCAAAAACATTTCTTCCCATTTGGCATAACATCCCATGTCTCTGC
ACAATCCTTCTTTGAAAATAAATATGGTAACTTAGATATATTTAGTACATTACGTTTCCTCTTG
CTTGATCGACATCATTCAAGAGCTCTTCAAAGCATTGTTCAGATCTTCAGTACTGGCCAGT
TTTCATACAGTCTCGGGGTTTTAAACCTTTGAAATCAAGGACACGACGTCTCCAGTCTACCTC
CGAGAGATTAGCTGAAACACAGAATATAGCGCCATCATTTCGTGAAGGGGTTTTCTTTGCGGGA
CAGAGGATCAGATGTTGAGAGTTTGGACAACTCATGAAAACCAAAAATATACCTGAAGCTCA
CCAAGATGCATTTAAAACTGGTTTTGCGGAAGGTTTTCTGAAAGCTCAAGCACTCACAAAAA
AACCAATGATTCCTTAAGGCGAACCCGTCTGATTCTCTTCGTTCTGCTGCTATTTCGGCATTTA
TGGACTTCTAAAAAACCCATTTTATCTGTCCGCTTCCGGACAACAACAGGGCTTAGTTCTGTC
AGTAGATCCTGTCCAGATGAAAAATGTCACCTTTGAACATGTTAAAGGGGTGGAGGAAGCTAA
ACAAGAATTACAGGAAGTTGTTGAATTCTTGAAAAATCCACAAAAATTTACTATTCTTGAGG
TAACTTCCAAAAGGAATTCTTTTAGTTGGACCCCCAGGGACTGGAAAGACACTTCTTGCCCG
AGCTGTGGCGGGAGAAGCTGATGTTCTTTTTATTATGCTTCTGGATCCGAATTTGATGAGAT
GTTTGTGGGTGTGGGAGCCAGCCGTATCAGAAATCTTTTAGGGAAGCAAAGGCGAATGCTCC
TTGTGTTATATTTATTGATGAATTAGATTCTGTTGGTGGGAAGAGAATTGAATCTCCAATGCA
TCCATATTCAAGGCAGACCATAAATCAACTTCTTGCTGAAATGGATGGTTTTAAACCCAATGA
AGGAGTTATCATAATAGGAGCCACAACTTCCCAGAGGCATTAGATAATGCCTTAATACGTCC
TGGTCTGTTTTGACATGCAAGTTACAGTTCCAAGGCCAGATGTAAAAGGTCGAACAGAAATTTT
GAAATGGTATCTCAATAAAATAAAGTTTGATCAATCCGTTGATCCAGAAATTATAGCTCGAGG
TACTGTTGGCTTTTCCGGAGCAGAGTTGGAGAATCTTGTGAACCAGGCTGCATTAAAAGCAGC
TGTTGATGGAAAAGAAATGGTTACCATGAAGGAGCTGGAGTTTTCCAAAGACAAAATTTCTAAT
GGGGCCTGAAAGAAGAAGTGTGGAAATTGATAACAAAAACAAAACCATCACAGCATATCATGA
ATCTGGTCATGCCATTATTGCATATTACACAAAAGATGCAATGCCTATCAACAAAGCTACAAT
CATGCCACGGGGGCCAACACTTGGACATGTGTCCCTGTTACCTGAGAATGACAGATGGAATGA
AACTAGAGCCCAGCTGCTTGCACAAATGGATGTTAGTATGGGAGGAAGAGTGGCAGAGGAGCT
TATATTTGGAACCGACCATATTACAACAGGTGCTTCCAGTGATTTTGATAATGCCACTAAAAT
AGCAAAGCGGATGGTTACCAAATTTGGAATGAGTGAAAAGCTTGGAGTTATGACCTACAGTGA
TACAGGGAAACTAAGTCCAGAAACCAATCTGCCATCGAACAAGAAATAAGAATCCTTCTAAG
GGACTCATATGAACGAGCAAAACATATCTTGAAAACTCATGCAAAGGAGCATAAGAATCTCGC
AGAAGCTTTATTGACCTATGAGACTTTGGATGCCAAAGAGATTCAAATGTTCTTGAGGGGAA
AAAGTTGGAAGTGAGATGATAACTCTCTTGATATGGATGCTTGCTGGTTTTATTGCAAGAATA
TAAGTAGCATTGCAGTAGTCTACTTTTACAACGCTTCCCCTCATTCTTGATGTGGTGTAATT
GAAGGGTGTGAAATGCTTTGTCAATCATTTGTACATTTATCCAGTTGGGTTATTCTCATT
TGACACCTATTGCAAAATTAGCATCCCATGGCAAATATATTTTGAAAAATAAAGAACTATCAG
GATTGAAAACAAAAA

96/550

FIGURE 96

MFSLSSTVQPQVTVPLSHLINAFTPKNTSVSLSGVSVSQNQHRDVVPEHEAPSSEPSLNLRD
LGLSELKIGQIDQLVENLLPGFCKGKNISSHWHTSHVSAQSFFENKYGNLDIFSTLRSSCLYR
HHSRALQSICSDLQYWPVFIQSRGFKTLKSRTRRLQSTSERLAETQNIAPSFVKGFLLRDRGS
DVESLDKLMKTKNIPEAHQDAFKTGFAEGFLKAQALTQKTNDLSLRTRLILFVLLLFGIYGLL
KNPFLSVRFRTTTTGLDSAVDPVQMKNVTFEHVKGVEEAKQELQEVVEFLKNPQKFTILGGKLP
KGILLVGPPGTGKTLLARAVAGEADVPFYYASGSEFDEMFGVGASRIRNLFREAKANAPCVI
FIDELDSVGGKRIESPMHPYSRQTINQLLAEMDGFKPNEGVIIGATNFPEALDNALIRPGRF
DMQVTVPRPDVKGRTEILKWYLNKIKFDQSVDPETIARGTVGFGAELENLVNQAALKAADV
KEMVTMKELEFSKDKILMGPERRSVEIDNKNKTITAYHESGHAIAYYTKDAMPINKATIMPR
GPTLGHVSLLPENDRWNETRAQLLAQMDVSMGGRVAEELIFGTDHITGASSDFDNATKIAKR
MVTKFGMSEKLGVMTYSDTGKLSPETQSAIEQEIRILLRDSYERAKHILKTHAKEHKNLAEAL
LTYETLDAKEIQIVLEGKKLEVR

Important features of the protein:**Transmembrane domain:**

amino acids 238-259

N-glycosylation sites.amino acids 28-32, 90-94, 230-234, 278-282, 535-539, 584-588,
623-627**N-myristoylation sites.**

amino acids 35-41, 266-272, 286-292, 325-331, 357-363, 599-605

Amidation site.

amino acids 387-393, 709-713

ATP/GTP-binding site motif A (P-loop).

amino acids 322-330

AAA-protein family proteins

amino acids 315-336, 343-386, 405-451

97/550

FIGURE 97

GATGGCGCAGCCACAGCTTCTGTGAGATTGATTTCTCCCCAGTCCCCTGTGGGTCTGAGGG
GACCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATCCCCAAA
ACAAGTTTTGACATTTCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTT
CCAGGCCTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACG
ACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAG
GCTTGGCAGTTTTTCTTACTCCTGTGGTCTCCAGATTTTCAGGCCTAAGATGAAAGCCTCTAGT
CTTGCCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAG
ACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTTCT
GAGATACGGGGCAGTGTGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACT
GAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTC
TATCTGGACAGGGTATTTAAAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGC
AGCCTCGCCAATTCCTTTCTTACCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACA
TGCCATTGTGGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTG
GAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGAACTAGACATTCTTCTGCAATGGATGGAG
GAGACAGAAATAGGAGGAAAGTGATGCTGCTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCT
TCAATACCTGCAGAGGAGGCATGACCCCAAACCACCATCTCTTTACTGTACTAGTCTTGTGCT
GGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCAATGATTGTCTTTATGCATCCCC
AATCTTAATTGAGACCATACTTGTATAAGATTTTTGTAATATCTTTCTGCTATTGGATATATT
TATTAGTTAATATATTTATTTATTTTTTGCTATTTAATGTATTTATTTTTTTTACTTGGACATG
AACTTTAAAAAATTCACAGATTATATTTATAACCTGACTAGAGCAGGTGATGTATTTTTAT
ACAGTAAAAAATAACCTTGTAATTTCTAGAAGAGTGGCTAGGGGGTTATTCATTTGTAT
TCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGATATTTGAAATTGAACCAATGAC
TACTTAGGATGGGTGTGGAATAAGTTTTGATGTGGAATTGCACATCTACCTTACAATTACTG
ACCATCCCCAGTAGACTCCCCAGTCCCATATTGTGTATCTTCCAGCCAGGAATCCTACACGG
CCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATACCAAAAAAAAAAAAAAAAAAAAAA

98/550

FIGURE 98

MKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNGFSEIRGSVQAKDGNIDIR
ILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKISSSLANSFLTIIKKDLRLC
HAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEETE

Signal sequence:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 107-110, 140-143

N-myristoylation site.

amino acids 51-56

Interleukin 10:

amino acids 9-176

99/550

FIGURE 99

[illegible]

100/550

FIGURE 100

MRLLPWFLLLLFGPWLLRKAVSAQIPESGRPQYLGLRPAAAGAGAPGQQLPPEPRSSDGLGVGR
AWSWAWPTNHTGALARAGAAGALPAQRTKRKPSIKAARAKKIFGWGDFYFRVHTLKFSLLVTG
KIVDHVNGTFSVYFRHNSSSLGNLSVSIVPPSKRVEFGGVWLPGPVPHPLQSTLALEGVLPGL
GPPLGMAAAAAGPGLGGSLLGGALAGPLGGALGVPGAKESRAFNCHVEYEKTNRARKHRPCLYD
PSQVCFTTEHTQSQAAWLCAKPFKVICIFVSFLSFDYKLVQKVC PDYNFQSEHPYFG

Important features of the protein:**Signal peptide:**

amino acids 1-22

Transmembrane domain:

amino acids 273-288

N-glycosylation sites.

amino acids 72-76, 133-137, 143-147, 149-153

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 93-97

N-myristoylation sites.amino acids 35-41, 58-64, 60-66, 81-87, 84-90, 184-190, 194-200,
203-209, 205-211, 206-212, 209-215, 217-223, 221-227, 224-230**Cytochrome b/b6 Qo site signature.**

amino acids 5-11

101/550

FIGURE 101

AATGCCCCATGCGCACCCACAGCTCGCGCTCCTGCAAGTGTTCTTTCTGGTGTTCCCCGATG
GCGTCCGGCCTCAGCCCTCTTCCTCCCCATCAGGGGCAGTGCCCACGTCTTTGGAGCTGCAGC
GAGGGACGGATGGCGGAACCCTCCAGTCCCCTTCAGAGGCGACTGCAACTCGCCCCGGCCGTGC
CTGGACTCCCTACAGTGGTCCCTACTCTCGTGACTCCCTCGGCCCCCTGGGAATAGGACTGTGG
ACCTCTTCCCAGTCTTACCGATCTGTGTCTGTGACTTGACTCCTGGAGCCTGCGATATAAATT
GCTGCTGCGACAGGGACTGCTATCTTCTCCATCCGAGGACAGTTTTCTCCTTCTGCCTTCCAG
GCAGCGTAAGGTCTTCAAGCTGGGTTTGTGTAGACAACCTCTGTTATCTTCAGGAGTAATTCCC
CGTTTCCTTCAAGAGTTTTTCATGGATTCTAATGGAATCAGGCAGTTTTGTGTCCATGTGAACA
ACTCAAACCTTAACTATTTCCAGAAGCTTCAAAGGTCAATGCAACCAACTTCCAGGCCCTGG
CTGCAGAGTTTGGAGGCGAATCATTTCACTTCAACATTCCAACTCAATCACCACCATCTTTTT
ACAGGGCTGGGGACCCCATTTACTTACTTCCCCAAGTGGTCTGTAATAAGCTTGCTGAGAC
AACCTGCAGGAGTTGGAGCTGGGGGACTCTGTGCTGAAAGCAATCCTGCAGGTTTCCTAGAGA
GTAAAGTACAACCTTGCACTCGTTTTTTCAAGAACCTGGCTAGTAGCTGTACCTTGGAATTCAG
CCCTCAATGCTGCCTCTTACTATAACTTCACAGTCTTAAAGGTTCCAAGAAGCATGACTGATC
CACAGAATATGGAGTTCAGGTTCTGTAACTTACCTCACAGGCTAATGCTCCTCTGTTGG
CTGGAAACACTTGTGAGAAATGTAGTTTCTCAGGTCACCTATGAGATAGAGACCAATGGGACTT
TTGGAATCCAGAAAGTTTCTGTGAGTTTGGGACAAACCAACCTGACTGTTGAGCCAGGCGCTT
CCTTACAGCAACACTTCATCCTTCGCTTCAGGGCTTTTCAACAGAGCACAGCTGCTTCTCTCA
CCAGTCCTAGAAGTGGGAATCCTGGCTATATAGTTGGGAAGCCACTCTTGGCTCTGACTGATG
ATATAAGTTACTCAATGACCCTCTTACAGAGCCAGGGTAATGGAAGTTGCTCTGTTAAAAGAC
ATGAAGTGCAAGTTTGGAGTGAATGCAATATCTGGATGCAAGCTCAGGTTGAAGAAGGCAGACT
GCAGCCACTTGCAAGCAGGAGATTTATCAGACTCTTCATGGAAGGCCAGACCAGAGTATGTTG
CCATCTTTGGTAATGCTGACCCAGCCCAGAAAGGAGGGTGGACCAGGATCCTCAACAGGCACT
GCAGCATTTCAAGCTATAAACTGTACTTCTGCTGTCTCATACCAGTTTCCCTGGAGATCCAGG
TATTGTGGGCATATGTAGGTCTCCTGTCCAACCCGCAAGCTCATGTATCAGGAGTTTCGATTCC
TATACCAGTGCCAGTCTATACAGGATTCTCAGCAAGTTACAGAAGTATCTTTGACAACCTCTTG
TGAACCTTTGTGGACATTACCCAGAAGCCACAGCCTCCAAGGGGCCAACCCTAAATGGACTGGA
AATGGCCATTGCACTTCTTTCCCTTCAAAGTGGCATTGAGCAGAGGAGTATTCTCTCAAAAAT
GCTCAGTCTCTCCATCCTTATCCTGTGCTCTTACTACTTGAGTTTCTAACCTAGAGACTA
TGTGAAGAAAAGAAAATAATCAGATTTTCAAGTTTTCCCTATGAGAACTCTGAGGCAGCCACTT
ATCTTGGCTAAATAGAACCCTCACCTGCTCATGACCAGAGAGCATTTAGGATAATAGATGACCT
AACTGAAGGAATCCTTGTATATGAAAGGAGTTATTTTAGAAAAGCAATAAAAATATTTTATTC
ATCNTAAAAAAAAA

102/550

FIGURE 102

MRT PQ LALLQVFFLVFPDGV RPQPSSSPSGAVPTSLELQRGTDGGTLQSPSEATATRPVPG L
PTVVPTLVTPSAPGNRTVDLFPVLPICVCDLTPGACDINCCCDRDCYLLHPRTVFSFCLPGSV
RSSSWVCVDNSVIFRSNSPFPSRVFMD SNGIRQFCVHVNN SNLNYFQKLQKVNATNFQALAAE
EGGESFTSTFQTQSPPSFYRAGDPILTYFPKWSVISLLRQPAGVGAGGLCAESNPAGFLESKS
TTCTRFFKNLASSCTLDSALNAASYNFTVLKVPRSM TDPQNMEFQVPVILTSQANAPLLAGN
TCQNVVSQVTYE IETNGTFGIQKVSVS LGQTNL TVEPGASLQQHFILRFRAFQQSTAASLTSP
RSGNPGYIVGKPLLALTDDISYSMTLLQSQGN GSCSVKRHEVQFGVNAISGCKLRLKKADCSH
LQQEIYQTLHGRPRPEYVAIFGNADPAQKGGWTRILNRHCSISAINCTSCCLIPVSLEIQVLW
AYVGLLSNPQAHVSGVRFLYQCQSIQDSQQVTEVSLTTLVNFVDITQKPQPPRGQPKMDWKWP
FDFFPFKVAFSRGVFSQKCSVSPILILCLLLLGVNLNLET M

Important features of the protein:**Signal peptide:**

amino acids 1-22

Transmembrane domains:

amino acids 484-505, 581-600

N-glycosylation sites.amino acids 78-82, 165-169, 179-185, 279-285, 331-337, 347-351,
410-414, 487-491**N-myristoylation sites.**

amino acids 30-36, 41-47, 124-130, 232-238, 236-242, 409-415

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 420-431

103/550

FIGURE 103

CCTAATTCTCAAGGTGATGCTATTTAGGAAGTCATAACTCATGTGAGTGGAGCCATGTGGGAT
TAAGAAGTGATAGGAGAGCTTGCTGTCTGTCTCTGCTCTCCACTGTGTGAGGATACAACAGGA
AGACAGCCATCTGGTGAGGAAGAGAGGGCCCTCGCCAGATACCGGACCTGCTGACACCTTGAT
CTTGGA CTCTCCATCTTCCAGGAAGGCCCTGACCTCAGTTGTTCCAGGGTAAAGAATTTGGGCA
GTGCCCACACCCACGCTGTTGGATAACATTTCTTCACCATAACCAGTGAGGGTGAATGTGTACA
CGCCCAGCTTCCTGCCTGTTACTCTCCACAGTATGCGAAGAATATCCCTGACTTCTAGCCCTG
TGCGCCTTCTTTTGTCTGCTGTTGCTACTAATAGCCTTGGAGATCATGGTTGGTGGTCACT
CTCTTTGCTTCAACTTCACTATAAAATCATTTGTCCAGACCTGGACAGCCCTGGTGTGAAGCGC
AGGTCTTCTTGAATAAAAAATCTTTTCCTTCAGTACAACAGTGACAACAACATGGTCAAACCTC
TGGGCCTCCTGGGGAAGAAGGTATATGCCACCAGCACTTGGGGAGAATTGACCCAAACGCTGG
GAGAAGTGGGGCGAGACCTCAGGATGCTCCTTTGTGACATCAAACCCAGATAAAGACCAGTG
ATCCTTCCACTCTGCAAGTCGAGATGTTTTGTCAACGTGAAGCAGAACGGTGCCTGGTGCAT
CCTGGCAGTTCGCCACCAATGGAGAGAAATCCCTCCTCTTTGACGCAATGAACATGACCTGGA
CAGTAATTAATCATGAAGCCAGTAAGATCAAGGAGACATGGAAGAAAGACAGAGGGCTGGAAA
AGTATTTTCAAGGAGCTCTCAAAGGGAGACTGCGATCACTGGCTCAGGGAATTCTTAGGGCACT
GGGAGGCAATGCCAGAACCGACAGGCAGAAGATCCACCTAGAGGTGATACCACGGCGGCGCAG
AGTTGTTTACCTGTGGTCTCGATCGCTGACAGCCTTGGCTCCCACTGCTGTGTGTTCCCTGA
GTCAAGTGGAGGCGGAGCCTGCAATGAGCGGAGATCGCGCCTCTGCATTCCAGTCTTGGCAAC
AGAGCAAGACTCCGTCTCAAAAAAAAAAATTTTTTTTTCAGTACATATTTTTTTAAAGATAGG
GCTGGGCACAGCAGCTCACATCTATAATCCCAACACTTTGGGAGGCCTAGGCAGGAGGATCAC
TTGAGCCCAGGAATCTGAAGCTGCAGTGAGCCTTTGCTCGTGAGATTGTGGACCTATGATCCT
ACCACCAGCCCACCTGGTTCTAACACCCCTCCTCTATGTGTGAGAGGGAGAGAAGAAAAGTG
AGGGAGAAAAGAGAGATAAGCAAAGAACAGAGAGGAAAAATGGAAAATAAGAGGAAATTGGGG
GAATTAAACAGAGGGGAGGGCATGGATCCCCGGGAGTTAGAAGAGTAGCAGCTTGTGGATTAC
TACGCAGTGGAGGAAGAAGAGTTGTTGGAAATTATTTGAGAGGTAGTATAATCATTTGTGAGG
CAGTTTTCTGCATTACCATTTCTCACAGACTAAGTTACTCATAAGCAAACGTGCAATTCACA
TTACACTGAAATTCTTCCCTAATACATCATTTGCATTGGAATAAAGTACGGTTTTCAAACAAC
CTGATATAGCAGAACTGACTGTATAAATTATGTGAGCACAGTGCAAGTAATCTTTGTTTGT
TGTTTGTTTTTTTGAGACAGAGTCTCACTCTATCTCCAGGCTGGAGTGTAGTGGTGCATCC
CGGCTCACTGCAACCTCGATCTCCAGGCTCAAGCGATTCCCCTGCCTCAGCCTCCTGAGTAG
CTGGGATTACAGGCATGAGCCACCACGCCCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGT
TTCACCCTGTTGGCCAGGCTGGTCTCGAACTACGGACCTCAGGTGATCTGCCCCCTCAGCCT
CTCAAAGTGCTGGGATTATAGCATGAGCCACTGAGCCCAGACACAAGTAGTTCTTTCTGATAA
ACACTTTAACTGAATGCA

104/550

FIGURE 104

MRRISLTSSPVRLLLFLLLLLIALEIMVGGHSLCFNFTIKSLSRPGQPWCEAQVFLNKNLFLQ
YNSDNNMVKPLGLLGKKVYATSTWGELTQTLGEVGRDLRMLLCDIKPQIKTSDPSTLQVEMFC
QREAERCTGASWQFATNGEKSLLFDAMNMTWTVINHEASKIKETWKKDRGLEKYFRKLSKGDC
DHWLREFLGHWPEAMPEPTGRRST

Important features of the protein:**Signal peptide:**

amino acids 1-23

Transmembrane domain:

amino acids 11-30 (possible type II protein)

N-glycosylation site.

amino acids 36-39, 154-157

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 2-5, 182-185, 209-212

Casein kinase II phosphorylation site.

amino acids 86-89, 93-96, 142-145, 185-188

N-myristoylation site.

amino acids 46-51

Amidation site.

amino acids 77-80, 207-210

105/550

FIGURE 105

TTTTCCGAGTGACCTTCTTG**ATG**CTGGCTGTTTCTCTCACCGTTCCTGCTTGGAGCCATGA
TGCTGCTGGAATCTCCTATAGATCCACAGCCTCTCAGCTTCAAAGAACCCCGCTCTTGCTTG
GTGTTCTGCATCCAAATACGAAGCTGCGACAGGCAGAAAGGCTGTTTGAAAATCAACTTGTTG
GACCGGAGTCCATAGCACATATTGGGGATGTGATGTTTACTGGGACAGCAGATGGCCGGGTCG
TAAAACTTGAAAATGGTGAAATAGAGACCATTGCCCGGTTTGGTTCGGGCCCTTGCAAAACCC
GAGATGATGAGCCTGTGTGTGGGAGACCCCTGGGTATCCGTGCAGGGCCCAATGGGACTCTCT
TTGTGGCCGATGCATACAAGGGACTATTTGAAGTAAATCCCTGGAAACGTGAAGTGAACTGC
TGCTGTCTCCGAGACACCCATTGAGGGGAAGAACATGTCCTTTGTGAATGATCTTACAGTCA
CTCAGGATGGGAGGAAGATTTATTTACCGATTCTAGCAGCAAATGGCAAAGACGAGACTACC
TGCTTCTGGTGATGGAGGGCACAGATGACGGGCGCCTGCTGGAGTATGATACTGTGACCAGGG
AAGTAAAAGTTTTATTGGACCAGCTGCGGTTCCCGAATGGAGTCCAGCTGTCTCCTGCAGAAG
ACTTTGTCCTGGTGGCAGAAACAACCATGGCCAGGATACGAAGAGTCTACGTTTCTGGCCTGA
TGAAGGGCGGGGCTGATCTGTTTGTGGAGAACATGCCTGGATTTCAGACAACATCCGGCCCCA
GCAGCTCTGGGGGGTACTGGGTGGGCATGTGACCATCCGCCCTAACCTGGGTTTTCCATGC
TGGATTTCTTATCTGAGAGACCCTGGATTAAAGGATGATTTTAAAGCTCTTTAGTCAAGAGA
CGGTGATGAAGTTTGTGCCGCGGTACAGCCTCGTCCTAGAACTCAGCGACAGCGGTGCCTTCC
GGAGAAGCCTGCATGATCCCGATGGGCTGGTGGCCACCTACATCAGCGAGGTGCACGAACACG
ATGGGCACCTGTACCTGGGCTCTTTAGGTCCCCCTTCCTCTGCAGACTCAGCCTCCAGGCTG
TT**TAG**CCCTCCAGATAGCTGCCCCCTGCCACGCAGGCCAGGAGTCTTCACACTCAGGCACCAG
GCCTGGTCCAGGAGGAGCTGTGGACACAGTCGTGGTTCAAGTGTCCACATGCACCTGTTAGTC
CCTGAGAGGTGGTGGGAATGGCTGCTTCATTCTCGAGGATGCCCGGGCCCCACCTGGGCTTG
TCTTTCTGTTTAGAGGGAAGTGTAACATATCTGCCATGAGGAACATAAATTCATGTAAAGCCA
TTTTCTCTTAAACAAAACAAAACCTTTCTAAGTACAATCATTCTCTAGGATTTGGGAAGCTCCT
TGCACTTGGAACAGGGCTCAGGTGGGTGGAGCAGTAAGGCACTACCCAGAGAGCTTGCTGCTG
CGGCCCTGTCCTGCGGCCTCAAAGTTCTTCTTTACTATATATAACGTGCGGTCATACCTTTCT
TCGTTGTGGTGGGGATGGAAGAGCAGAGGGAGCATGGCCCAGGGGTGTTGAGGCCAGCGGTGA
GAGCCGTGTTAGCCAAGACATGGAACGTGTGTTCTCAAGGGTTATGTGGGGCGTGGGCTCTCCA
TAGTGTGTATGAAAAGCTTGTTGACTCTAGCGGCTCAGAGAGGACTTTGCTGGGTTTCTTTCT
GTGAATATCTCCGTGCTGACCATGCTGGAATTGGATGATTCTGCAATTCGGGACCTACTGCAG
GGGTCCGTTTAGTAACGTCTTGTCTGTGATCTTTGTTCTTGACCTCTAGACCCCAAGATGTGA
ACAGTGCACGTGTTAATGTCATCTTTGCTCATGTGTTATAAGCCCCAAGTTGCTGTATATTTT
CACAAGTATGTCTACACACTGG

106/550

FIGURE 106

MLAVSLTVPLL GAMMLLES PIDPQPLSFKEPPLL LGVLHPNTKLRQAERLFENQLVGPESIAH
IGDVMFTGTADGRVVKLENGEIETIARFGSGPCKTRDDEPVCGRPLGIRAGPNGTLFVADAYK
GLFEVNPWKREVKLLLSSETPIEGKNMSFVNDLTVTQDGRKIYFTDSSSKWQRRDYLLLVMEG
TDDGRLLEYDTVTREVKVLLDQLRFPNGVQLSPAEDFVLVAETTMARIRRVYVSGLMKGGADL
FVENMPGFDPNIRPSSSGGYWVGMSTIRPNPGFSMLDFLSERPWIKRMIFKLFSQETVMKFVP
RYSLVLELSDSGAFRRSLHDPDGLVATYISEVHEHDGHLVYLGSRSPFLCRLSLQAV

Important features of the protein:**Signal peptide:**

amino acids 1-13

Transmembrane domain:

amino acids 1-21 (possible type II)

N-glycosylation sites.

amino acids 116-119, 152-155

Casein kinase II phosphorylation sites.

amino acids 19-22, 27-30, 98-101, 146-149, 221-224, 286-289, 332-335

N-myristoylation sites.

amino acids 71-76, 92-97, 189-194, 244-249, 338-343

Amidation site.

amino acids 164-167

107/550

FIGURE 107

AACGAAGCGTGCGCGCTTTGGTAACCGGCTAGAAATCCCGCACGCGCGCCTGCCTCCTCTCCC
CAGGCCTGAGCTGCCCCCTCCCACTGCCTTTCTTCTTCCCGCGAGTCAGAAGCTTCGCGAGGG
CCCAGAGAGGCGGTGGGGTGGGCGACCCCTACGCCAGCTCCGGGCGGGAGAAAGCCCACCCTCT
CCCGCGCCCCCAGGAAACCGCCGGCGTTCGGCGCTGCGCAGAGCCATGGAATTCTCCTGGCTGG
AGACGCGCTGGGCGCGGCCCTTTTACCTGGCGTTCGTGTTCTGCCTGGCCCTGGGGCTGCTGC
AGGCCATTAAGCTGTACCTGCGGAGGCAGCGGCTGCTGCGGGACCTGCGCCCCCTTCCCAGCGC
CCCCACCCACTGGTTCTTGGGCACCAGAAAGTTTATTTCAGGATGATAACATGGAGAAGCTTG
AGGAAATTATTGAAAAATACCCCTCGTGCCTTCCCTTTCTGGATTGGGGCCCTTTCAGGCATTTT
TCTGTATCTATGACCCAGACTATGCAAAGACACTTCTGAGCAGAACAGATCCCAAGTCCCAGT
ACCTGCAGAAATCTCACCTCCACTTCTTGGAAAAGGACTAGCGGCTCTAGACGGACCCCAAGT
GGTTCCAGCATCGTCGCCTACTAACTCCTGGATTCCATTTTAAACATCCTGAAAGCATACATTG
AGGTGATGGCTCATTCTGTGAAAATGATGCTGGATAAGTGGGAGAAGATTTGCAGCACTCAGG
ACACAAGCGTGAGGTCTATGAGCACATCAACTCGATGTCTCTGGATATAATCATGAAATGCG
CTTTCAGCAAGGAGACCAACTGCCAGACAAACAGCACCCATGATCCTTATGCAAAAGCCATAT
TTGAACTCAGCAAAATCATATTTACCGCTTGTACAGTTTGTGTATCACAGTGACATAATTT
TCAAACCTCAGCCCTCAGGGCTACCGCTTCCAGAAGTTAAGCCGAGTGTGTAATCAGTACACAG
ATACAATAATCCAGGAAAGAAAGAAATCCCTCCAGGCTGGGGTAAAGCAGGATAACACTCCGA
AGAGGAAGTACCAGGATTTTCTGGATATTGTCTTTCTGCCAAGGATGAAAGTGGTAGCAGCT
TCTCAGATATTGATGTACACTCTGAAGTGAGCACATTCTGTGGCAGGACATGACACCTTGG
CAGCAAGCATCTCCTGGATCCTTTACTGCCTGGCTCTGAACCCTGAGCATCAAGAGAGATGCC
GGGAGGAGGTACAGGGGCATCCTGGGGGATGGGTCTTCTATCACTGGGACCAGCTGGGTGAGA
TGTCGTACACCACAATGTGCATCAAGGAGACGTGCCGATTGATTCCTGCAGTCCCGTCCATTT
CCAGAGATCTCAGCAAGCCACTTACCTTCCCAGATGGATGCACATTGCCTGCAGGGATCACCG
TGGTCTTAGTATTTGGGGTCTTCACCACAACCCTGCTGTCTGGAAAAACCCAAAGGTCTTTG
ACCCCTTGAGGTCTCTCAGGAGAATTCTGATCAGAGACACCCCTATGCCTACTTACCATTCT
CAGCTGGATCAAGGAACTGCATTGGGCAGGAGTTTGCCATGATTGAGTTAAAGGTAACCATTG
CCTTGATTCTGCTCCACTTCAGAGTGACTCCAGACCCCAAGGCTCTTACTTTCCCAACC
ATTTTATCCTCAAGCCCAAGAATGGGATGTATTTGCACCTGAAGAACTCTCTGAATGTTAGA
TCTCAGGGTACAATGATTAAACGTACTTTGTTTTTCGAAGTTAAATTTACAGCTAATGATCCA
AGCAGATAGAAAGGATCAATGTATGGTGGGAGGATTGGAGGTTGGTGGGATAGGGGTCTCTG
TGAAGAGATCCAAAATCATTTCTAGGTACACAGTGTGTAGCTAGATCTGTTTCTATATAACT
TTGGGAGATTTTCAGATCTTTTCTGTAAACTTTCACTACTATTAATGCTGTATACACCAATA
GACTTTCATATATTTTCTGTTGTTTTTAAATAGTTTTTCAGAATTATGCAAGTAATAAGTGCA
TGTATGCTCACTGTCAAAAATTCCTAACACTAGAAAATCATGTAGAATAAAAATTTTAAATCT
CACTTCACTTAGCCGACATTCCATGCCCTGACCAATCCTACTGCTTTTCTAAAAACAGAATA
ATTTGGTGTGCATTCTTTCAGACTTTTTCCTATACATTTTATATGTAGAAATGTAGCAATGTA
TTTGTATAGATGTGATCATTCCCTATATTGTTATTGATTTTTTCACTTAATAAAAATTCACCT
TATTCCTTAAAA

108/550

FIGURE 108

MEFSWLETRWARPFYLAFFVCLALGLLQAIKLYLRRQRLRLRPFPPAPPTHWFLGHQKFIQD
DNMEKLEEIIIEKYPRAPFPWIGPFQAFFCIYDPDYAKTLLSRTDPKSQYLQKFSPELLGKGLA
ALDGPKWQHRRLTTPGFHFNILKAYIEVMAHSVKMMLDKWEKICSTQDTSVEVYEHINSMSL
DIIMKCAFSKETNCQTNSTHDPYAKAIFELSKII FHRLYSLLYHSDIIFKLSPQGYRFQKLSR
VLNQYTDTTIIQERKKSLQAGVKQDNTPKRKYQDFLDIVLSAKDESGSSFSDIDVHSEVSTFLL
AGHDTLAASISWILYCLALNPEHQERCREEVRGILGDGSSITWDQLGEMSYTTMCIKETCRLI
PAVPSISRDL SKPLTFPDGCTLPAGITVVLSIWGLHHNPAVWKNPKVFDPLRFSQENSQRHP
YAYLPFSAGSRNCIGQEFAMIELKVTIALILLHFRVTPDPTRPLTFPNHFILKPKNGMYLHLK
KLSEC

Important features of the protein:**Signal peptide:**

amino acids 1-29

Transmembrane domains:

amino acids 310-330, 397-413, 459-473

N-glycosylation site.

amino acids 206-210

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 265-269, 504-520

N-myristoylation sites.

amino acids 25-31, 298-304, 353-359, 450-456, 456-462

Cytochrome P450 cysteine heme-iron ligand signature:

amino acids 447-457

Cytochrome P450 cysteine heme-iron ligand proteins.

amino acids 444-475

109/550

FIGURE 109

GGCGTTCGGGCTCAACTTTGGCGTCGTGAGATTCTTGTGAGGCGTCTGCCTGGAAGCCGGC
AGCAATTTTGCTTCTTTAAAGAGAAAAAGAAGGCTAGGGACTCAGATTCCTGGATTCTGAGAT
CCAGACCAGCTCCTCCCAGACCTCTCCAGAAGAAGCCATGGGAACCCCTCGTATCCAGCATTT
GCTGATCCTCCTGGTCCTAGGAGCCTCCCTCCTGACCTCGGGCCTAGAGCTGTATTGTCAAAA
GGGTCTGTCCATGACTGTGGAAGCAGATCCAGCCAATATGTTTAACTGGACCACAGAGGAAGT
GGAGACTTGTGACAAAGGGGCACTTTGCCAGGAAACCATACTAATAATTAAAGCAGGGACTGA
GACAGCCATTTTGGCCACGAAGGGCTGCATCCCGGAAGGGGAGGAGGCCATAACAATTGTCCA
GCACTCTTCACCTCCCGGCCTGATCGTGACCTCCTACAGTAACTACTGTGAGGATTCCTTCTG
TAATGACAAAGACAGCCTGTCTCAGTTTTGGGAGTTCAGTGAGACCACAGCTTCCACTGTGTC
AACAACCCTCCATTGTCCAACCTGTGTGGCTTTGGGGACCTGTTTCAGTGCTCCTTCTCTTCC
CTGTCCCAATGGTACAACCTCGATGCTATCAAGGAAAACCTTGAGATCACTGGAGGTGGCATTGA
GTCGTCTGTGGAGGTCAAAGGCTGTACAGCCATGATTGGCTGCAGGCTGATGTCTGGAATCTT
AGCAGTAGGACCCATGTTTGTGAGGGAAGCGTGCCACATCAGCTGCTCAACCTCGAAA
GACTGAAAATGGGGCCACCTGTCTTCCCATTCCTGTTTGGGGGTTACAGCTACTGCTGCCATT
GCTGCTGCCATCATTATTCACTTTTCCTAGAAGGCATTCTGGGCCTGGGTCTGAGGACAT
CTTTTTTGACTGGGAGCCTTCTTACTGTTGAGGTTCAACAAGCTGAGGAGTAGATGGGAATTT
GAGGGAGAATACAGAGATACTATGAACGTATTTGACATTTTAAACAATTTCTGCTATAATT
TTTGTATGCAGTAGGCGTTACTAATAAACATTTCTGCTGTGA

110/550

FIGURE 110

MGTPRIQHLLILLVLGASLLTSGLELYCQKGLSMTVEADPANMFNWTTEEVETCDKGALCQET
ILIIKAGTETAILATKGCIPEGEEAITIVQHSSPPGLIVTSYSNYCEDSFCDKDSLSQFWEF
SETTASTVSTTLHCPTCVALGTCFSAPSLPCPNGTTRCYQGKLEITGGGIESSVEVKGCTAMI
GCRLMSGILAVGPMFVREACPHQLLTQPRKTENGATCLPIPVWGLQLLLPLLLPSFIHFS

Important features of the protein:**Signal peptide:**

amino acids 1-23

Transmembrane domain:

amino acids 184-201

N-glycosylation sites.

amino acids 45-49, 159-163

N-myristoylation sites.amino acids 31-37, 70-76, 99-105, 147-153, 160-166, 174-180,
175-181

111/550

FIGURE 111

CGAGAAGAGGACAGAGGAGACTGAGCAAAGGGGGGTGGGCTCCAGGCGACCCCTAGCCCAATTCTGCCCCCTCCAT
CCCAAGGGGCAGAGAAATTGTCTTTTGTGCTGACTCCTACGAGGAAAAAAAAAAAAAAAAAAAAAACCATTAA
AGGGAAAGATAAACGGAGACGGAGGAAAGGTGGCAGCCAGATTACTTAGAGAGGCACAGAGGAGAGAGATCGGGG
TGAGTCGCCATGGGGACTCCCAGGGGCCAGCACCCCGCCCTCCCGAGCTGCTGTTCCCTAATTCTGCTGAGCTGT
CCCTGGATCCAGGGTCTGCCCCGAAGGAGGAGGAGATATTGCCAGAGCCTGGAAGTGAGACCCACCGGTGGCC
TCTGAGGCCCTGGCTGAAGTGTTCATGGGGCCCTGCTGAGGAGGGGCCAGAGATGGGCTACCTGCCAGGATCT
GATCCGGACCCACGCTAGCCACCCCTCCGGCCGGCCAGACTCTCGCAGTGCCCTCCCTGCCACGGGGCCACTGAG
CCGGGGACAGGGCCTCTGACAACAGCCGTACCCCTAACGGGGTACAGGGGGCAGGCCCCACTGCGCCAGAACTG
CTGACCCCGCCCCAGGAACCCACAGCCCCACCCACCCAGCCCTGCCCTCCCGAGGGCTCCCTTGGGCTGAG
GGAGGAGAGGAGGAGACGACGACCACCATCATCACGACAACCTGTTACCACTACGGTGACCAGCCAGTTCTG
TGTAATAACACATCTCCGAGGGCGAAGGGTATGTGGAGTCTCCAGATCTGGGGAGCCCCGTACGCCGACCCCTG
GGGCTCCTGGACTGCACTTACAGCATCCATGTCTACCTGGCTACGGCATTGAGATCCAGGTGCAGACGCTGAAC
CTGTACAGGAAGAGGAGCTCCTGGTGTGGCTGGTGGGGGATCCCCAGGCCTGGCCCCCGACTCCTGGCCAACT
TCATCCATGCTTGGAGAAGGACAAGTCTTTCGGAGGCCCAACCAACCGGCTGCTTCTGCACTTCCAGAGCCCCACG
GTCCCAAGGGCGGTGGCTTCAGGATCCACTATCAGGCCTACCTCCTGAGCTGTGGCTTCCCTCCCGCGCGGCC
CATGGGGACGTGAGTGTGACGGACCTGCACCTGGGGGCACTGCCACCTTCACTGTGATTCGGGCTACCAGCTG
CAGGGAGAGGAGACCCTCATCTGCCTCAATGGCACCCGGCCATCCTGGAACGGTGAAACCCCCAGCTGCATGGCA
TCCTGTGGTGGCACCATCCACAATGCCACCTGGGCCGCATCGTGTCCCAGAGCCTGGGGGAGCCGTAGGGCCC
AACCTCACCTGCCGTTGGGTCAATTGAAGCAGCTGAGGGGCGCCGGCTGCACCTGCACCTTGAAGGGTCTCGCTG
GATGAGGACAATGACCGGCTGATGGTGGCTCAGGGGGCAGCCCCCTATCCCCCGTGATCTATGATTCGGACATG
GACGATGTCCCCGAGCGGGGTCTCATCAGTGACGCCAGTCCCTCTACGTGGAGCTGCTGTGAGAGACACCTGCC
AATCCCTGCTGTGAAGCCTTCGATTTGAAGCCTTTGAGGAGGATCGCTGCTTCGCCCTTCCCTGGCACATGGA
AATGTCACTACCACGGACCTGAGTATCGCCAGGGGCACTGGCAACCTTCTCGTGCCTCCAGGATATGCCCTG
GAGCCCCCTGGGCCCCCAATGCCATCGAATGTGTGGATCCACAGAACCCCACTGGAACGACACAGAGCCGGCC
TGCAAGCCATGTGTGGAGGGGAGCTGTCCGAACAGCTGGCGTGGTCTCTCTCCCGACTGGCCCCAGAGCTAT
AGCCCGGCCAAGACTGCGTGTGGGGCGTGACGTCCAGGAAGAGAAGCGCATCTTGCTCCAAGTTGAGATATTG
AATGTGCGGGAAGGGGACATGCTGACGCTGTTCGACGGGGACGGTCCCAGCGCCCGAGTCTTGGCCCCAGCTGCGG
GGACCTCAGCCGCGCCCGCCCTTCTCTCTCTGGGCCCCGACCTCACACTGCAGTTTCAGGCACCGCCCGGGCCC
CCAAATCCAGGCCCTGGGCCAGGGCTTCGTATTGCACTTCAAAGAGGTCCCGAGGAACGACAGTGCCCCGAGCTG
CCACCTCCGGAGTGGGGCTGGAGAACGGCATCCACGGGGACCTGATCCGGGGCACGGTGTCTACCTACCAGTGC
GAGCCTGGCTACGAGCTGTAGGCTCCGACATTTCACTTGCCAGTGGGACCTGTCTGGAGCGCCGCGCCGCC
GCCTGCCAAAAGATCATGACTTGTGCTGACCCTGGCGAGATTGCCAACGGGCACCGCACCGCCTCGGACGCCGGC
TTCCCCGTTGGCTCCCACGTCCAGTACCGCTGCCTGCCAGGGTACAGCCTCGAGGGGGCAGCCATGCTCACCTGC
TACAGCCGGGACACAGGCACACCCAAGTGAGCGGATAGGGTCCCCAAATGCGCCTTGAAGTACGAGCCGTGCCTG
AACCCGGGGGTTCCCGAGAATGGCTACCAGACGCTGTACAAGCACCCTACCAGGCGGGCGAGTCTCTGCGCTTC
TTCTGTATGAGGGCTTTGAGCTTATCGGCGAGGTCAACATCACCTGTGTGCCCGGCCACCCCTCCAGTGGACC
AGCCAGCCCCACTCTGCAAAGTGACCCAGACCACAGATCCATCACGGCAGCTGGAAGGGGGGAACCTGGCCCTG
GCCATCCTGCTGCCTTAGGCTTGGTCATTGTCTCGGAGTGGCGTTTACATCTACTACCAAGCTTCAGGGA
AAGTCCCTTTTCGGGCTTCTCGGGCTCCCACTCTACAGCCCCATCACCGTGGAGTCGGACTTCAGCAACCCGCTG
TATGAAGCTGGGGATACGCGGGAGTATGAAGTTTCCATCTGAACCCCAAGACTACAGCTGCAGGACCCAGGACGC
CCCTCCCTCCTCATTCGGGCAGAGGGAAATACGGGACCCGGTCTGTGCCTCCTGGCTGCCCTCCTCCCTGGCTG
TGTAATAGTCTCCCTATCCACGAGGGGGCTTTGATGGCCCTGGAGATCCTACAGTAAATAAACCAGCATCCTG
CCGCCAAAAAA

112/550

FIGURE 112

MGTPRAQHPPPPQLLFLILLSCPWIOGLPLKEEEILPEPGSETPTVASEALAE LLHGALLRRG
PEMGYLPGSDPDPTLATPPAGQTLAVPSLPRATEPGTGPLTTAVTPNGVRGAGPTAPELLTPP
PGTTAPPPPPSPASPGPPLGPEGGEEETTTTIIITTTT VTTTSPVLCNNNISEGEGYVESPD
GSPVSR TLGLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELLVLAGGSGPGLAPRLLANSSMLG
EGQVLRSP TNRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFPPRPAHGDVSVTDLHPGGTATFH
CDSGYQLQGEETLICLN GTRPSWNGETPSCMASCGGTIHNATLGRIVSPEPGGAVGPNLTCRW
VIEAAEGRRLHLHFERVSLDEDNDRLMVRSGGSPLSPVIYDSMDDDVPERGLISDAQSLYVEL
LSETPANPLLLSLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPGALATFSCLPGYALEPPGPPN
AIECVDPTTEPHWNDTEPACKAMCGGELSEPAGVVLSPDW PQSYSPGQDCVWGVHVQEEKRILL
QVEILNVREGDMLTLFDGDGPSARVLAQLRGPQPRRLLSSGPD LTLQFQAPPGPPNPGLGQG
FVLHFKEVPRNDTCPELPPPEWGWRTASHGDLIRGTVLTYQCEPGYELLGSDILTCQWDLWS
AAPPACQKIMTCADPGEIANGHRTASDAGFPVGS HVQYRCLPGYSLEGAAMLTCYSRDTGTPK
WSDRVPKCALKYEPCLNPGVPENGYQTLYKH HYQAGESLRFFCYEGFELIGEVTITCVPGHPS
QWTSQPPLCKVTQTTPSRQLEGGNLALAILLPLGLVIVLGS GVIYYTKLQGKSLFGFSGSH
SYSPITVESDFSNPLYEAGDTREYEVSI

Important features of the protein:**Signal peptide:**

amino acids 1-27

Transmembrane domain:

amino acids 842-864

N-glycosylation sites.amino acids 176-180, 222-226, 247-251, 332-336, 355-359, 373-377,
473-477, 517-521, 641-645**Tyrosine kinase phosphorylation site.**

amino acids 61-69

N-myristoylation sites.amino acids 2-8, 84-90, 111-117, 114-120, 190-196, 198-204,
235-241, 309-315, 333-339, 351-357, 472-478, 484-490, 528-534,
626-632, 665-671, 775-781, 842-848**Amidation site.**

amino acids 384-388

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 12-23

CUB domain proteins profile.

amino acids 202-218, 376-392, 553-569

113/550

FIGURE 113

GCCGCGGGCGGAGCTGCCTGCCGGTCCGCGCGCGCGCTCCGCACTCCTCGGCCCTCGGGCGGTGCATGGGACGG
GGCGCCGCGGAGCAGGAGCGCGCCCGTCCGGGGTGCTCGGGCCGCGCGGGAGCCCACTGTGGGGCTCGGGCATG
GCGGGCCGAGGACCTGAGCTCTCCTCAGGGGAGCGGGGAGGCAGCTGCTGGCCGCGCATGGGGACGGAGTGGGG
CCGTGCGCGCGCGCGCGAGCCGTGAGCGCCGAGCCACCGCCGCGCTACCTCAGCCCTTCGCGAAGCGCCGGGCA
GCTCGGGAACATGGCCCTGGAGCGGCTCTGCTCGGTCTCTCAAAGTGTGTTAATAACAGTACTGGTAGTGGAAGG
GATTGCCGTGGCCCAAAAAACCAAGATGGACAAAATATTGGAATCAAGCATATTCTGCAACCCAGTGTGGCAT
TTGGGTTTGAACACAGCAATGGAGGTCATTTTGCTTCGCCAAATTATCCTGACTCATATCCACCAAACAAGGAGTG
TATCTACATTTTGAAGCTGCTCCACGTCAAAGAATAGAGTTGACCTTTGATGAACATTATTATATAGAACCATC
ATTTGAGTGTGCGTTTGATCACTTGAAGTTCGAGATGGGCCATTGGTTTCTCTCTCTTATAGATCGTTACTG
TGGCGTGAAAAGCCCTCCATTAATTAGATCAACAGGGAGATTCAATGTGGATTAAAGTTTAGTTCTGATGAAGAGCT
TGAAGGACTGGGATTTGAGCAAAAATATTCATTTATTCAGATCCAGACTTTACTTACCTAGGAGGTATTTTAA
TCCATTCCAGATTGTCTAGTTCGAGCTCTCGGGAGCTGATGGAATAGTGCCTCTAGTCAGGTAGAACAAGAGGA
GAAAACAAAACAGGCCAAGCCGTTGATTGCATCTGGACCATTAAAGCCACTCCAAAAGCTAAGATTATTTGAG
GTTCTAGATTATCAATGGAGCACTCAATGAATGCAAGAGAACTTCGTTGCAGTCTATGATGGAAGCAGTTC
TATTGAAAACCTGAAGGCCAAGTTTTCAGCACTGTGGCCAATGATGTAATGCTTAAACAGGAATTGGAGTGAT
TCGAATGTGGGCAGATGAAGGTAGTCGGCTTAGCAGGTTTCGAATGCTCTTACTTCTCTTGTGGAGCCTCCCTG
CACAAGCAGCACTTTCTTTGCCATAGCAACATGTGCATCAATAATTCTTTAGTCTGTAATGGTGTCAAAATTG
TGCATACCCTTGGGATGAAAATCATTGTAAAGAAAAGAAAAAGCAGGAGTATTTGAACAAATCACTAAGACTCA
TGGAACAATTATTGGCATTACTTCAGGATTGTCTTGGTCTTCTCATTATTTCTATTTTAGTACAAGTGAAACA
GCCTCGAAAAAAGGTCAATGGCTTGCAAAACCGCTTTTAAATAAAACCGGGTTCCAAGAAGTGTGATCCTCTCA
TTATGAAGTGTCTTCACTAAGGGACAAGAGATTTCTGCAGACCTGGCAGACTTGTGGAAGAATTGGACAACCTA
CCAGAAGATGCGGCGTCTCCACCGCTCCCGCTGCATCCACGACCACCTGTGGGTGCGAGGCCCTCCAGCGT
CAAACAAAGCAGGACCAACCTCAGTTCCATGGAACCTCTTTCGAAATGACTTTGCACAACCACAGCCAATGAA
AACATTTAATAGCACCTTCAAGAAAAGTAGTTACACTTTCAAACAGGGACATGAGTGCCTTGAGCAGGCCCTGGA
AGACCGAGTAATGGAGGAGATTCCCTGTGAAATTTATGTCAGGGGGCGAGAAGATTCTGCACAAGCATCCATATC
CATTGACTTCTAATCTCTGCTAATGGTGATGTGAATCTTAGGGTGTGTACGTACGCAGCCTCCAGGGCACCAT
ACTGTTTCCAGCAGCCAACCCCTTTCTCCCATCACAACTACGAAGACCTTGATTACCGTTAACCTATTGTATGG
TGATGTTTTTATTCTCTCAGGCAGTCTATATATGTTAAACCAATCAAGGAACCTTACTCTATTCAAGTGAACAAT
AATCATCTCTATTGCTTGGTGTCTTTATAGGAAGCACTGCCAGTTAAAGAGCATTAGAAGAGGTGGTTGGATGG
AGCCAGGCTCAGGCTGCCTCTTCGTTTGTAGCAACAAGAAGACTGCTCTTGACTGATAACAGCTCTGTCAATATTT
TGATGCCACAATAAACTTGATTTTTTTTACATTCTTTTATTTTCTCTTCTCTAAATTTAATTTGTTTTATAA
GCCTATCGTTTTTACCATTTCATTTTCTACATAAGTACAAGTGTTAATGTACCACATACTTCAGTATAGGCATT
TGTTCTTGAGTGTGTCAAATAACAGCTAGTTACTGTGCCAATTAAGACCCAGTTGTATTTACCCATCTGTTTCT
TCTTGGCTAATCTCTGACTTCTGCCTTTTAAATTAAGTGGGCCCTTATTCCTTATTTTCTGTGAGAAATAATAGAT
GATATGATTTATTACCTTTCAATTATATTTTCTCAGTTATACTAGAAAATTTCAATCCTGGGATATATGTAC
CATTGTGAGCTATGACTAAAAATTTGAAAAAGATAAAAATTTCTAGCAAGCCTTTGAAGTTTACCAAGTATAGTC
ACATTCACTGACAGCCATTTCATTCCAGTAAAGAATCATTTCATTCACTTTGGGAGAGGCCCTATAATTACATTTA
TTTGCAATGTTTCTCTTCGCTAGATTGTACATAGCTCCCATTTCTGTTGGTTTTGCTTACAGCATATGGTAACCA
AGGTTAGATGCCAGTTAAATTCCTTAGAAATTGGATGAGCCTTGAGATTGCTTCTTAAGTGGGACATGACATTT
TTCTAGCTCTTATCAAGAATAACAACCTCCACTTTTTTTTAACTGCACTTTTGACTTTTTTTATGGTATAAAAA
CAATAATTTATAAACATAAAAGCTCATTGTGTTTTTTAGACTTTTGATATTATTTGATACTGTACAACTTTATT
AAATCAAGATGAAAGACCTACAGGACAGATTCTTTTCACTGTTCACATCAGTGGCTTTGTATGCAAAATAGCTGT
GTTGGACCTGGACGCTATAACTTATTGTAAAGACCTTGGAATGTGGACATAAGCTCTTCTTCTTCTTTGTTAC
TGTATTTAGTTTGTGATAAATTTTCACTGTGTGATATTTATGCTCTAATCACTACACAAATCCCATATTAAAA
TATACATTGTACCTGAAAAAAA

114/550

FIGURE 114

MALERLCSVLKVLITVLVVEGIAVAQKTQDGQNIGIKHIPATQCGIWVRTSNGGHFASPNYP
DSYPPNKECIYILEAAPRQRIELTFDEHYIIEPSFECRFDHLEVRDGPFGFSPLIDRYCGVKS
PPLIRSTGRFMWIKFSSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIPDCQFELSGADGIVR
SSQVEQEEKTKPGQAVDCIWTIKATPKAKIYLRFLDYQMEHSNECKRNFVAVYDGSSSIENLK
AKFCSTVANDVMLKTGIGVIRMWADEGSRLSRFRMLFTSFVEPPCTSSTFFCHSNMCINNSLV
CNGVQNCAYPWDENHCKEKKKAGVFEQITKTHGTIIGITSGIVLVLLIISILVQVKQPRKKVM
ACKTAFNKTGFQEVFDPPHYELFSLRDKEISADLADLSEELDNYQKMRRSSTASRCIHDHHC
SQASSVKQSRTNLSSMELPFRNDFAPQPMKTFNSTFKKSSYTFKQGHECPEQALEDRVMEEI
PCEIYVRGRED SAQASISIDF

Important features of the protein:**Signal peptide:**

amino acids 1-22

Transmembrane domain:

amino acids 348-369

N-glycosylation sites.

amino acids 311-315, 385-389, 453-457, 475-479

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 426-430, 479-483

N-myristoylation sites.amino acids 22-28, 32-38, 54-60, 186-192, 279-285, 318-324,
348-354, 352-358, 441-447

115/550

FIGURE 115

GGTCTCTGTCTTGGCTGTGGCTCCTGCGCTCTGGCTGAGCCATGTTTCCTTCTCCTCGCCCTC
CTCACTGAGCTTGGAAGACTGCAAGCCCACGAAGGTTCTGAAGGAATATTTCTGCATGTCACA
GTTCCACGGAAGATTAAGTCAAATGACAGTGAAGTTTCAGAGAGGAAGATGATTTACATCATT
ACAATTGATGGACAACCTTACACTCTACATCTCGGAAAACAATCATTCTTACCCCAGAACTTT
TTGGTTTATACATATAATGAACTGGATCTTTGCATTCTGTGTCTCCATATTTTATGATGCAT
TGCCATTACCAAGGATATGCTGCCGAATTTCCAAATTCATTTGTGACACTCAGTATATGTTCT
GGTCTCAGGGGATTTCTCCAGTTTGAAAATATCAGTTATGGAATTGAACCAGTAGAATCTTCA
GCAAGATTTGAGCATATAATTTATCAAATGAAAAATAATGATCCAAATGTATCCATTTTAGCA
GTAAATTACAGTCATATTTGGCAGAAAGACCAGCCCTACAAAGTTCCTTTAACTCACAGATA
AAAAATCTTTCAAAATATTACCCCAATATCTGGAAATATACATTATAGTGGAAGAAAGCTTTG
ATGTTTACCCAGTTCAAATTGACTGTTATACTGTCTTCCTTGGAAATTGTGGTCAAATGAAAC
CAGATTTCACCAAGTGGGGATGCTGATGATATATTACAAAGATTTTGGCATGGAAACGGGAC
TATCTCATCCTACGGCCCATGACATAGCATACTTACTTGTTTACAGGAAACATCCTAAATAT
GTGGGAGCAACATTTCTGGCACCAGTATGCAATAAAAGCTATGATGCAGGTATTGCTATGAT
CCAGATGCAATAGGTTTGGAGGGATTTTCGGTTATTATAGCTCAACTGCTTGGCCTTAATGTA
GGATTAACATATGATGACATCACTCAGTGTCTGTCTGAGAGCTACATGCATCATGAATCAT
GAAGCAGTGAGTGCCAGTGGTAGAAAGATTTTGTAGCAACTGCAGCATGCACGACTATAGATAT
TTTGTTCAAAATTTGAGACTAAATGCCTTCAGAAGCTTTCAAATTTGCAACCATTACATCAA
AATCAACCAGTGTGTGGTAATGGGATTTTGGAAATCCAATGAAGAATGTGACTGTGGTAATAAA
AATGAATGTCAATTTAAGAAGTGCTGTGATTATAACACATGTAACTGAAGGGCTCAGTAAAA
TGTGGTTCTGGACCATGTTGTACATCAAAGTGTGAGTTGTCAATAGCAGGCACTCCATGTAGA
AAGAGTATTGATCCAGAGTGTGATTTTACAGAGTACTGCAATGGAACCTCTAGTAATTGTGTT
CCTGACACTTATGCACTGAATGGCCGTTTGTGCAAGTTGGGAACTGCCTATTGCTATAACGGA
CAATGTCAAACACTGATAACCAGTGTGCCAAGATATTTGGAAAAGGTGCTCAAGGTGCTCCA
TTTGCCTGTTTTAAAGAAGTTAATTCTCTGCATGAAAGATCTGAAAACCTGTGGTTTTAAAAAT
TCACAACCATTACCTTGTGAACGGAAGGATGTTCTCTGTGGAAAATTAGCTTGTGTTTCAGCCA
CATAAAAATGCTAATAAAAGTGACGCTCAATCTACAGTTTATTTCATATATTCAAGACCATGTA
TGTGTATCTATAGCCACTGGTTCCCTCCATGAGATCAGATGGAACAGACAATGCCTATGTGGCT
GATGGCACCATGTGTGGTCCAGAAATGTACTGTGTAAATAAAACCTGCAGAAAAGTTCAATTA
ATGGGATATAACTGTAATGCCACCACAAAATGCAAAGGGAAAGGGATATGTAATAATTTTGGT
AATTGTCAATGCTTCCCTGGACATAGACCTCCAGATTGTAAATTCCAGTTTGGTTCCCCAGGG
GGTAGTATTGATGATGGAAATTTTCAGAAATCTGGTGACTTTTATACTGAAAAGGCTACAAT
ACACACTGGAACAACTGGTTTATTCTGAGTTTCTGCATTTTCTGCCGTTTTTCATAGTTTTTC
ACCACTGTGATCTTTAAAGAAATGAAATAAGTAAATCATGTAACAGAGAGAATGCAGAGTAT
AATCGTAATTCATCCGTTGTATCAGAAAGCGATGACGTGGGACATTAATATTGCACAGAACCT
CCATAGCAAATAACCTAAAGGAACGAATGTGCTTTATTTATAACCTTACGTTATCCCCAATGC
ATTGTAAATGTCAAACTTTTGGAAAATAAAGCCTGCGTGCCCTCCC

116/550

FIGURE 116

MFLLALLTELGRQLQAHEGSEGI FLHVTVPRKIKSNDSEVSEKMIYIITIDGQPYTLHLGKQ
SFLPQNFLVYTYNETGSLHSVSPYFMMHCHYQGYAAEFNSFVTLSSICSGLRGFLQFENISYG
IEPVESARFEHIIYQMKNNDPNVSILAVNYSHIWQKDQPYKVPLNSQIKNLSKLLPQYLEIY
IIVEKALMFTQFKLTVILSSLELWSNENQISTSGDADDILQREFLAWKRDYLILRPHDIAYLLV
YRKHPKYVGATFPGTVCNKSYDAGIAMYPDAIGLEGFSVIIAQLLGLNVGLTYDDITQCFCLR
ATCIMNHEAVSASGRKIFSNCSMHDYRYFVSKFETKCLQKLSNLQPLHQNPVCGNGILESNE
ECDCGNKNECQFKKCCDYNTCKLKGSVKCGSGPCCTSKCELSIAGTPCRKSIDPECDFTEYCN
GTSSNCVPTDYALNGRLCKLGTAYCYNGQCQTDTNQCAKIFGKGAQGAPFACFKEVNSLHERS
ENCGFKNSQPLPCERKDVLCGLACVQPHKNANKSDAQSTVYSYIQDHVCVSIATGSSMRSDG
TDNAYVADGTMCGPEMYCVNKTCKRVHLMGYNCNATTKCKGKGICNNFGNCQCFPGHRPPDCK
FQFGSPGGSIDGNGFQKSGDFYTEKGYNTHWNNWFILSFCIFLPFFIVFTTVIFKRNEISKSC
NRENAEYNRNSSVVSESDDVGH

Important features of the protein:**Signal peptide:**

amino acids 1-16

Transmembrane domain:

amino acids 665-684

N-glycosylation sites.amino acids 36-39, 76-79, 122-125, 149-152, 156-159, 177-180,
270-273, 335-338, 441-444, 537-540, 587-590, 601-604, 703-706**Casein kinase II phosphorylation sites.**amino acids 74-77, 208-211, 221-224, 304-307, 337-340, 346-349,
376-380, 415-418, 499-502, 639-642, 708-711**Tyrosine kinase phosphorylation site.**

amino acids 243-249

N-myristoylation sites.amino acids 53-58, 79-84, 266-271, 298-303, 372-377, 403-408,
408-413, 442-447, 462-467, 469-474, 488-493, 567-572, 610-615,
616-621, 634-639**Amidation site.**

amino acids 328-331

117/550

FIGURE 117

CCCACGCGTCCGCGGACGCGTGGGGCTCAGTGGGCGTCGCGCGAAGGCTAAGGGAGTGTGGCG
GGCGGCTCCGGGAGCCAACATGCCTCGGTATGCGCAGCTGGTCATGGGCCCCGCGGGCAGCGG
GAAGAGCACCTACTGTGCCACCATGGTCCAGCACTGTGAAGCCCTCAACCGGTCTGTCCAAGT
TGTAACCTGGATCCAGCAGCAGAACACTTCAACTACTCCGTGATGGCTGACATCCGGGAAGT
GATCGAGGTGGATGATGTAATGGAGGATGATTCTCTGCGATTCCGGTCCCAACGGAGGATTGGT
ATTTTGCATGGAGTACTTTGCCAATAATTTTGA CTGGCTGGAGAACTGTCTTGGCCATGTAGA
GGACGACTATATCCTTTTTGATTGTCCAGGTCAGATTGAGTTGTACACTCACCTGCCTGTGAT
GAAACATCTGGTCCAGCAGCTCGAGCAGTGGGAGTCCGAGTCTGTGGAGTTTTTCTTGTTGA
TTCTCAGTTCATGGTGGAGTCATTCAAGTTTATTTCTGGCATCTTGGCAGCCCTGAGTGCCAT
GATCTCTCTAGAAATCCGCAAGTCAACATCATGACAAAAATGGATCTGCTGAGTAAAAAAGC
AAAAAAGGAAATTGAGAAATTTTTAGATCCAGACATGTATTCTTTATTAGAAGATTCTACAAG
TGACTTAAGAAGCAAAAAATTCAAGAACTGACTAAAGCTATATGTGGACTGATTGATGACTA
CAGCATGGTTTCGATTTTTACCTTACGATCAGTCAGATGAAGAAAGCATGAACATTGTATTGCA
GCATATTGATTTTGCCATTCAATATGGAGAAGACCTAGAATTTAAAGAACCAAAGGAACGTGA
AGATGAGTCTTCCTCTATGTTTGACGAATATTTTCAAGAATGCCAGGATGAATGAAGAGTTTA
CTAAAAGTAACCATCTAAAGAGCTTGTGGCCAAACCAGCAGAACATTCTTCTCTTCAAAGGAT
GCAATAGTAGAAAGCTACTTATTTTAATGAAAAAAGTAAACTTCGTTCTTTATCAGCCTCA
TGCCTGAATCAAATTTTAAATTATTCTGAACTGCTGCTGTTTAAAGTGGAATCTTTTAGTAT
TATAACAGCATCACTTTAGATTTTGTAAGTCAAATGAAATGAATGCACATAGATTTATATA
TAAATTAGCACCTGAGCTAAGGTAAAGGCCGGTCTAACTTATTTTCACTTTTTGTATTATTT
TTGAGATGCAGGAATTACTGTAACAAAATATGTATGTCCGAAGGGAAAAAGCTGCAAGGATAT
ATATAAGACCACTGCTTATCTGTATCTTCCATTTTCCTATATTGAAAATGTATATTATTTAT
ATAACTTAAAAAGTAAAAATAACTATGTTTTGAGAT

118/550

FIGURE 118

MPRYAQLVMGPAGSGKSTYCATMVQHCEALNRSVQVVNLDPAAEHFNYSVMADIRELIEVDDV
MEDDSLRFPGNGGLVFCMEYFANNFDWLENCLGHVEDDYILFDCPGQIELYTHLPVMKHLVQQ
LEQWEFRVCGVFLVDSQFMVESFKFISGILAALSAMISLEIPQVNIMTKMDLLSKKAKKEIEK
FLDPDMYSLLEDSTSDLRSKKFKKLTKAICGLIDDYSMVRFLPYDQSDEESMNIVLQHIDFAI
QYGEDLEFKEPKEREDESSSMFDEYFQECQDE

Important features of the protein:**Signal peptide:**

amino acids 1-29

Transmembrane domain:

amino acids 151-170

N-glycosylation sites.

amino acids 31-35, 47-51

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 212-216

Tyrosine kinase phosphorylation site.

amino acids 189-197

N-myristoylation sites.

amino acids 13-19, 76-82, 154-160

ATP/GTP-binding site motif A (P-loop).

amino acids 10-18

119/550

FIGURE 119

GGGCGCTGGGAGACACCGGACGCCCGCTCGGCTGCGCTGCGGCTCAGGCCCCCGCTCGGGCCC
GACCCGCTCGGTACACGCCGGCTCGGGCGCGCACCTGCCGGCTGCGGCCCCAGGGCCATGCGG
AGGCCACAGAGGAGGCCGGCGGCCACGCGCATCCCGTAGCCAGGTGGCCCAGGTCTGCACCG
CGGCGGCCTCGGCGCC**ATG**GAGCCCCCGTATTGCTGACGGCGCACTACGATGAGTTCCAAGA
GGTCAAGTACGTGAGCCGCTGCGGCGCGGGGGGCGCGCGGGGGCCTCCCTGCCCCGGGCTT
CCCGTTGGGCGCTGCGCGCAGCGTACCGGGGCCCGGTCCGGGCTGCCGCGCTGGAACCGGCG
CGAGGTGTGCCTGCTGTGCGGGCTGGTGTTCGCCGCCGGCCTCTGCGCCATTCTGGCGGCTAT
GCTGGCCCTCAAGTACCTGGGCCCCGGTCGCGGCCGGCGGCGGCGCCTGTCCCGAGGGCTGCC
TGAGCGCAAGGCCTTCGCGCGCGCCGCTCGCTTCCTGGCCGCCAACCTGGACGCCAGCATCGA
CCCATGCCAGGACTTCTACTCGTTCGCCTGCGGCGGTTGGCTGCGGCGCCACGCCATCCCCGA
CGACAAGCTCACCTATGGCACCATCGCGGCCATCGGCGAGCAAAACGAGGAGCGCCTACGGCG
CCTGCTGGCGCGGCCCGGGGGTGGGCTGGCGGCGCGGCCAGCGCAAGGTGCGCGCCTTCTT
CCGCTCGTGCTCGACATGCGCGAGATCGAGCGACTGGGCCCCGCGACCCATGCTAGAGGTCTAT
CGAGGACTGCGGGGGCTGGGACCTGGGCGGCGCGGAGGAGCGTCCGGGGGTGCGGCGCGCATG
GGACCTCAACCGGCTGCTGTACAAGGCGCAGGGCGTGTACAGCGCCGCCGCGCTCTTCTCGCT
CACGGTCAGCCTGGACGACAGGAACCTCCTCGCGCTACGTATCCGCATTGACCAGGATGGGCT
CACCTGCCAGAGAGGACCTGTACCTCGCTCAGGATGAGGACAGTGAGAAGATCCTGGCAGC
ATACAGGGTGTTCATGGAGCGAGTGCTCAGCCTCCTGGGTGCAGACGCTGTGGAACAGAAGGC
CCAAGAGATCCTGCAAGTGGAGCAGCAGCTGGCCAACATCACTGTGTGAGATATGACGACCT
ACGGCGAGATGTGAGTCCATGTACAACAAGGTGACGCTGGGGCAGCTGCAGAAGATCACCCC
CCACTTGGCGGTGGAAGTGGCTGCTAGACCAGATCTTCAGGAGGACTTCTCAGAGGAAGAGGA
GGTGGTGCTGCTGGCGACAGACTACATGCAGCAGGTGTGCGAGCTCATCCGCTCCACACCCCA
CCGGGTCTGCACAACCTACCTGGTGTGGCGCGTGGTGGTGCTCTGAGTGAACACCTGTCCCC
GCCATTCCGTGAGGCACTGCACGAGCTGGCACAGGAGATGGAGGGCAGCGACAAGCCACAGGA
GCTGGCCCCGGGTCTGCTTGGGCCAGGCCAATCGCCACTTTGGCATGGCGCTTGGCGCCCTCTT
TGACATGAGCACTTCTCAGCCGCCAGCAAAGCCAAGGTGCAGCAGCTAGTGGAAGACATCAA
GTACATCCTGGGCCAGCGCCTGGAGGAGCTGGACTGGATGGACGCCGAGACCAGGGCTGCTGC
TCGGGCCAAGCTCCAGTACATGATGGTGATGGTGGCTACCCGGACTTCTGCTGAAACCCGA
TGCTGTGGACAAGGAGTATGAGTTTGAGGTCCATGAGAAGACCTACTTCAAGAACATCTTGAA
CAGCATCCCCTTCAGCATCCAGCTCTCAGTTAAGAAGATTCCGCAGGAGGTGGACAAGTCCAC
GTGGCTGCTCCCCCACAGGCGCTCAATGCCTACTATCTACCCAACAAGAACCAGATGGTGT
CCCCGCGGGCATCCTGCAGCCCACCCTGTACGACCCTGACTTCCCACAGTCTCTCAACTACGG
GGGCATCGGCACCATCATTGGACATGAGCTGACCCACGGCTACGACGACTGGGGGGGCCAGTA
TGACCGCTCAGGGAACCTGCTGCACTGGTGGACGGAGGCCTCCTACAGCCGCTTCTGCGAAA
GGCTGAGTGCATCGTCCGTCTCTATGACAACCTTCACTGTCTACAACCAGCGGGTGAACGGGAA
ACACACGCTTGGGGGAGAACATCGCAGATATGGGCGTCTCAAGCTGGCCTACCACGCCTATCA
GAAGTGGGTGCGGGAGCACGGCCCAGAGCACCCACTTCCCCGGCTCAAGTACACACATGACCA
GCTCTTCTTCAATTGCCTTTGCCCAGAACTGGTGCATCAAGCGGCGGTGCGAGTCCATCTACCT
GCAGGTGCTGACTGACAAGCATGCCCCTGAGCACTACAGGGTGTGGGCAAGTGTGTCCAGTT
TGAGGAGTTTGGCCGGGCTTTCCACTGTCCCAAGGACTCACCCATGAACCCTGCCCCACAAGTG
TTCCGTGTGG**TGA**GCCTGGCTGCCCCGCTGCACGCCCCCACTGCCCCCGCACGAATCACCTCC
TGCTGGCTACCGGGGCAGGCATGCACCCGGTGCCAGCCCCGCTCTGGGCACCACCTGCCTTCC
AGCCCCCTCAGGACCCGGTCCCCCTGCTGCCCCCTCACTTCAGGAGGGGCCTGGAGCAGGGTGA
GGCTGGACTTTGGGGGGCTGTGAGGGAAATATACTGGGGTCCCCAGATTCTGCTCTAAGGGGG
CCAGACCCTCTGCCAGGCTGGATTGTACGGGCCCCACCTTCGCTGTGTTCTTGCTGCAAAGTC
TGGTCAATAAATCACTGCACTGTTAAAAA

120/550

FIGURE 120

MEPPYSLTAHYDEFQEVKYVSRGAGGARGASLPPGFPLGAARSVTGARSGLPRWNRREVCLL
SGLVFAAGLCAILAAMLALKYLGPVAAGGGACPEGCPERKAFARAARFLAANLDASIDPCQDF
YSFACGGWLRRAIPDDKLTGTIAAIGEQNEERLRLLARPGGGPGGAAQRKVRAFFRSCLD
MREIERLGPRPMLEVIEDCGGWDLGGAERPGVAARWDLNRLLYKAQGVYSAAALFSLTVSLD
DRNSSRYVIRIDQDGLTLPERTLYLAQDEDESEKILAAAYRVFMERVLSLLGADAVEQKAQEILQ
VEQQLANITVSEYDDLRRDVSSMYNKVTLGQLQKITPHLRWKWLLDQIFQEDFSEEEVVLLA
TDYMQQVSQILIRSTPHRVLHNYLVWRVVVVLSEHLSPPFFREALHELAQEMEGSDKPQELARVC
LGQANRHFGMALGALFVHEHFSAAASKAKVQQLVEDIKYILGQRLEELDWMDAETRAAARAKLQ
YMMVMVGYPDFLLKPDADVDEKEYEFEVHEKTYFKNILNSIPFSIQLSVKKIRQEVDKSTWLLPP
QALNAYYLPKNQMVFAPAGILQPTLYDPDFPQSLNYGGIGTIIGHELTHGYDDWGGQYDRSGN
LLHWWTEASYSRFLRKAECIVRLYDNFTVYNQVRVNGKHTLGENIADMGVCLKLAYHAYQKWVRE
HGPEHPLPRLKYTHDQLFFIAFAQNWCIKRRSQSIYLQVLTDKHAPEHYRVLGVSVSQFEFEGFR
AFHCPKDSMPMNAHKCSVW

Important features of the protein:**Transmembrane domain:**

amino acids 64-88

N-glycosylation sites.

amino acids 255-259, 322-326, 656-660

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 722-726

N-myristoylation site.amino acids 24-30, 26-32, 27-33, 40-46, 47-53, 65-71, 148-154,
169-175, 170-176, 237-243, 450-456, 604-610, 607-613**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 85-96

Prenyl group binding site.

amino acids 772-777

Neutral zinc metalloproteinases, zinc-binding region signature.

amino acids 609-619

Neutral zinc metalloproteinases, zinc-binding region proteins.

amino acids 609-619

121/550

FIGURE 121

CGGACTGCCCCGACCGCGCGATGGAGTCGACCGGCAGCGTCGGGGAGGCCCCGGGCGGACCCC
GGGTGCTGGTGGTGGGCGGCGGCATCGCGGGGCTGGGCGCGGCAGAGGCTCTGCGGCCACT
CCGCCTTCCCGCACCTGCGGGTCCTGGAGGCCACGGCCCGCGCCGGGGGCCGCATCCGCTCGG
AGCGCTGCTTCGGTGGCGTGGTGGAGGTGGGCGCGCACTGGATCCATGGGCCCTCCCGGGGTA
ACCCCGTCTTCCAGCTGGCTGCTGAGTACGGGCTGCTGGGGGAGAAGGAGCTGTCCCAGGAGA
ACCAGCTGGTGGAGACCGGGGGTCACGTGGGCCTGCCCTCCGTGAGCTACGCCAGCTCCGGGG
CCAGCGTGAGCCTCCAGCTGGTGGCGGAGATGGCGACTCTGTTCTACGGCCTGATAGACCAGA
CCCGGGAGTTCTTGCACGCTGCAGAGACCCCGGTGCCAGCGTCGGGGAGTACCTCAAGAAGG
AGATTGGCCAGCACGTGGCCGGCTGGACAGAGGATGAGGAGACCAGGAAGCTGAAGCTGGCCG
TCCTGAACTCCTTCTTCAACCTGGAATGCTGTGTGAGCGGCACCCACAGCATGGACCTGGTGG
CCCTGGCACCCCTTTGGGGAGTATACCGTGCTGCCGGGGCTGGACTGCACCTTTTCTAAGGGCT
ATCAAGGACTCACAACTGCATGATGGCCGCCCTGCCGGAGGACACTGTAGTTTTTGTAGAAGC
CTGTGAAGACCATCCACTGGAACGGGTCTTCCAGGAGGCAGCCTTTCCCGGGGAGACCTTTC
CAGTGTCCGTAGAGTGTGAGGATGGAGACCGGTTCCCGGCGCACCATGTCATCGTCACCGTGC
CCTTAGGTTTTCTTAGGGAACATTTGGACACCTTCTTTGACCCTCCCCTGCCGGCTGAGAAGG
CAGAAGCAATCAGGAAGATAGGCTTTGGGACCAACAACAAAATCTCCTGGAGTTTGAGGAGC
CCTTCTGGGAGCCAGACTGCCAGCTGATCCAGCTGGTGTGGGAGGACACGTCGCCCCCTGGAGG
ATGCTGCCCCCTGAGCTACAGGACGCTGGTTCGGGAAGCTCATTGGCTTTGTGGTCTGCTG
CCTTTGCGTCTGTCCACGTTCTCTGTGGGTTCAATTGCCGGAAGTTGAGTCTGAGTTCATGGAGA
CTCTGTCCGATGAAGAAGTACTTCTGTGTCTCACCCAAGTGCTCCGGAGAGTGACAGGAAACC
CACGGCTCCCCGCGCCCAAGAGCGTCCTGCGGTCTCGCTGGCACAGCGCCCCGTACACTAGGG
GGTCTACAGCTACGTGGCCGTGGGCAGTACTGGGGGCGACCTGGACCTGCTGGCTCAGCCCC
TCCCTGCAGACGGCGCCGGCGCCAGCTCCAGATCCTGTTTGCGGGGGAAGCCACACATCGCA
CGTTTTACTCCACGACGCACGGGGCTCTGCTGTCCGGATGGAGGGAGGCCGACCGCCTCCTCA
GTCTGTGGGCCCCGCAGGTGCAGCAGCCAGGCCGAGGCTCTAGCTGGGCCCAGCCTACTCTG
TTCCACCCGTGTCGGGGGTAGGCTGGGACCGTCATTTCTTCTGACAGATTTTCAGTCTGGCTTG
AAATTTGGGGATGTTAATGAGGGTCTCTGGTTTTTTGGTAACCAGGGCCACCTTCTCAGTTCT
TGTGTCTGTTATTGGAGTCTGGCCAGGGTTGACTTGAGCTGAGACACCAGATGCTCACGGAGA
TGCTGGACACATAAAGCAAGTTACAGCCACAAAAA

122/550

FIGURE 122

MESTGSGEAPGGPRVLVVGGGIAGLGAAQRLCGHSAPHLRVLEATARAGGRIRSERCFGGV
VEVGAHWIHGPSRGNPVFQLAAEYGLLGEKELSQENQLVETGGHVGLPSVSYASSGASVSLQL
VAEMATLFYGLIDQTREFLHAAETPVPSVGEYLLKKEIGQHVAGWTEDEETRKLKLAVLNSFFN
LECCVSGTHSMDLVALAPFGEYTVLPGLDCTFSKGYQGLTNCMAALPEDTVVFEKPVKTIHW
NGSFQEAAFPGETFPVSVECEDGDRFPAHHVIVTVPLGFLREHLDTFDPPLPAEKAEAIRKI
GFGTNNKIFLEFEEFPWEPCQLIQLVWEDTSPLEDAAPELQDAWFRKLIGFVVLPAFASVHV
LCGFIAGLESEFMETLSDEEVLLCLTQVLRRTGNPRLPAPKSVLRSRWHSAPYTRGSYSYVA
VGSTGGDLDLLAQPLPADGAGAQLQILFAGEATHRTFYSTTHGALLSGWREADRLLSLWAPQV
QQPRPRL

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 364-385

N-glycosylation site.

amino acids 253-257

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 408-412

N-myristoylation sites.amino acids 20-26, 21-27, 25-31, 105-111, 119-125, 164-170,
216-222, 227-233, 443-449, 484-490**Aminooxidase Flavin containing amine oxidase:**

amino acids 23-497

123/550

FIGURE 123

CGGACGCGTGGGGGAAGATGGATAAATAATTCTGTACACGTGCCCTGGCCTCTGGAGCTCAGCTGCCAGTCCAC
GTCTAGGGAATCTTAGCATCTGGGACCAAGACACTTTACAGCAATCATCACCCTTTGCAGAGGAGGTGAGCTCAC
CAGGACTCATCTGCCATTTAGACACTTTTGTGCTACCTGCCAGGTGGCCCCACTGCTGACGAGAGATGTTGGGA
TCTCTCAGTCTCCCCGACTCCTTGAAGCCAGTATCGCTGACCAGTCTTGTCTTCTCATGCACCTCCTCCT
CCTTCAGCCTGGGGAGCCGAGCTCAGAGGTCAAGGTGCTAGGCCCTGAGTATCCCATCCTGGCCCTCGTCGGGGA
GGAGGTGGAGTTCCCGTGCCACCTATGGCCACAGCTGGATGCCAGCAAATGGAGATCCGCTGGTTCCGGAGTCA
GACCTTCAATGTGGTACACCTGTACCAGGAGCAGCAGGAGCTCCCTGGCAGGCAGATGCCGGCGTTCCGGAACAG
GACCAAGTTGGTCAAGGACGACATCGCCTATGGCAGCGTGGTCTGCAGCTTACAGCATCATCCCCTCTGACAA
GGGCACATATGGCTGCCGCTTCCACTCCGACAACCTTCTCTGGCGAAGCTCTCTGGGAACCTGGAGGTAGCAGGGCT
GGGCTCAGACCCTCACCTCTCCCTTGAGGGCTTCAAGGAAGGAGGCATTAGCTGAGGCTCAGATCCAGTGGCTG
GTACCCCAAGCCTAAGGTTAGTGGAGAGACCACAGGGACAGTGCCTGCCTCCAGAGTTTGAAGCCATCGTCTG
GGATGCCCAGGACCTGTTAGTCTGGAACATCTGTGGTTGTCCGAGCGGGAGCCCTCAGCAATGTGTCCGCTCTC
CATCCAGAATCTCCTCTTGAGCCAGAAGAAAGATTGGTGGTCCAGATAGCAGACGTGTTCTGATCCCGGAGCTC
TGCGTGGAAAGAGCGCGTTCTGTGCGACCCCTGCCGCTGCTGTTGGTCTCGCGGCGCTGGCGCTGGGCGTCTCCG
GAAGCAGCGGAGAAGCCGAGAAAAGCTGAGGAAGCAGGCGGAGAAGAGACAAGAGAACTCACTGCAGAGCTGGA
AAAGCTTCAGACAGAGCTTGACTGGAGACGGGCTGAAGGCCAGGCTGAGTGGAGAGCAGCCAAAAATATGAGT
GGATGTGACGCTGGACCCGGCCTCGGCGCACCCAGCCTGGAGGTGTGGAGGATGGCAAGAGCGTGTCTCCCG
CGGGCGCGCCAGGCCCGCGCCTGGCCACCCGACGCGTTCTGGAGCAGACGTGCGCGCTGAGCCTGGAGCG
GTTCTCCGCGCGCCGCTACTGGGAGGTGACGTGGGCGCGCCGAGCCGCTGGTCTCGGCGCCTGCCTGGC
CGCGGTGCCGCGCGGGGCTGCGCGCCTGAGCCCTGCGGCGGCTACTGGGTGCTGGGGCTGTGGAACGGCTG
CGAGTACTTCTGCTCGGCCCGCACCGCGCTCGCGCTCACCCTGCGCGTGCCTCCGCGCGCTGGGCGTCTTCT
GGACTACGAGGCCGGAGAGCTGTCTTCTTCAACGTGTCCGACGGCTCCACATCTTACCTTCCACGACACCTT
CTCGGGCGCGCTCTGTGCGTACTTCAGGCCCAGGGCCAGCAGCGCGCGCAACATCCGATCCCTGACCATCTG
CCCCGCTGCCGCTTAGAGGGACGGGCTCCCCGAGAGAAAGGAGACAGTACACCTGGCTACAGCCCTATGAGCCGC
GGACCCCGCCTGGACTGGTGGTGGAGGCGCCCTCGTGCCGCGGGACTGGCCCCGGGGGGCCCCCTGGATCCAG
GCCAGCCTTTGCTCTCCTGCTCCGCTCTGAAGGGAGCAGGTGCCAGCCAAAATGTGAGCGAGGGGGACAAAGA
GAGGGACCTTTGCCTACGTAGATGTGTATGTGTAGTGCGATTTTCTTCAAGGAAGGAGACAAGTCCAAGCTCG
TTTGTGATTGTGGGACTGAGCGAAGGAGTACAAATATATCCAGTCTGCTCAGAGCTGGGGTGCTCACGGTGGG
GGTGGGCAAGAAGCCAGCATGGAAGAAAGAGGGGAGAAAACCTTGGTGACTGCCTTAGAGGGATCAGTTAATTTG
TATAGTTTTATATTTTTTGTATATGTTTGTAGCTCTAAAAAGGTGAGATGCAATAACACTTCGTAAGCAACGA
GTTACCTAAGTAAGGCTCAGATCCTAGTTTTAAAAACCATTTCCCATTAATGAAGTTGGAGGAACAGCTGCT
TCTGAGCCGGGGCAAAAATTTCAAGGTGAGCCTGGAGCATTTGTGTGGTGAAGTAAAAATAAGGCTCAAAACGT
GACGGCAACCCGGCAAAAGGTTAGGGAGCCAGGCCGAAGGGGCTCACTGACCAATTGTGGGACAATTTGAACAT
CAGGATGAATAATGACAGGAGAGATTATAACACACTGAATAAAAAACATAATCCATGAGTTTCTGCTGATACTCAA
ATTTCTTTTAAAAAGGAGAAAACAGGAAGGTTCTTTTGGAGGTGAAATCTAATTATTGGTGAGAGTCTTGAGA
ACAGGCTGTTTCCAGTCTCAAAGCAGTAACCTTATACACTACTTATAAGTTTGAAGGGGAAAGGTTACCTTTAC
AATGGAGACATCTACCAGATCATCCAAGTGATTAAATTTAATCATCAATGATGGGACCAAGGACATTATTAGT
TTGACAACTGGGGAAAGAAGTGTCTTCAACCCCTACCCCAAGACATTCTCTGTGCGCCAGGCTGGAGTGCA
GCCTCAACCTCCTGGGCCAAGTGATCCTCCACCTCAGCACACAACCATGCCCAATTTAAGTGCCTTATAG
AGACGGGGTCTCACTTTGTTACCCAGGCTGGTCTCAAACCTCTGCGCTCAAGCAATCCTCCACCTGGGCTCC
CAAAATGCTGGGTGTACAGGCATGAGCCGCTGTGCTGGCTTCATTTTCAAGTGTGACATTTGTACTGTGGCTA
TGTAGGAGAACATTTCTTCTTAGCAACATACTGAAGTTTTTAGATATTAATTACCACAGTGTCTGCCACTGA
ATTTCCAGTGACTAAGTGGAAAAATATAAAACATATGAATATAAAGAAAGAAAGAGACAAGTCAAATGTAGTAAA
ATGACAACACTTGGTGACTCTAGGTGACTGGTCGACAGATGTTTCTGACTATCAATGTGGCTTTGCTGTGGGT
TTGAAATTTTCAAACCTAAGAGTTGGGTGGCGGGGAGAAGGATACCAAAAACTAAGTGATTATCTTTGGATG
GGAAATGTTTGGTAATTGCATTCTTAAATGTCTTCTTGTATTTTTAATGTTCAATAATGTATATGTATCAG
TTCTGTAATAAAGGGGAAAAACACTTTTCA

124/550

FIGURE 124

MVDLSVSPDSLKPVSLTSSLVFLMHLLLLQPGEPSSEVKVLGPEYPILALVGEEVEFPCHLWP
QLDAQQMEIRWFRSQTFNVVHLYQEQQELPGRQMPAFRNRTKLVKDDIAYGSVVLQLHSIIPS
DKGTYGCRFHSDFNFSGEALWELEVAGLGSDPHLSLEGFKEGGIQLRLRSSGWYPKPKVQWRDH
QGQCLPPEFEAIVWDAQDLFSLETSVVVRAGALSNVSVSIQNLLLSQKKELVVQIADVFPVGA
SAWKSAFVATLPLLLVLAALALGVLRKQRRSREKLRKQAEKRQEKLTAELEKLQTELDWRRAE
GQAEWRAAQKYAVDVTLDPASAHPSLEVSSEDGKSVSSRGAPPGPAPGHPQRFSEQTCALSLE
FSAGRHYWEVHVGRRSRWFLGACLAAPRAGPARLSPAAGYWVLGLWNGCEYFVLAPHRVALT
LRVPPRRLGVFLDYEAGELSFFNVSDGSHIFTFHDTFSGALCAYFRPRAHDGGEHPDPLTICP
LPVRGTGVPEENDSDTWLQPYEPADPALDWW

Important features of the protein:**Signal peptide:**

amino acids 1-34

Transmembrane domain:

amino acids 247-272

N-glycosylation sites.

amino acids 102-106, 139-143, 224-228, 464-468, 516-520

Tyrosine kinase phosphorylation site.

amino acids 105-114

N-myristoylation sites.

amino acids 129-135, 220-226, 399-405, 423-429, 480-486

Amidation site.

amino acids 390-394

125/550

FIGURE 125

TATAGTCCCAGCTACTCATGGGGCTGATGCAGGTTGAGGCAGGAGGTTTCATGAGCCCAGGAGGTTGGAGCTGTAA
TGAGCTAGGATTCTGCCTCTGCACTCCTAGCTGGATGACAGAGCAAGACCCTGTCTCAAAAAAGAAAAA
AAAAAGAATGCATGAACCAGACATGACAGTTCTTGGCCTCAAAGATCTTCCAAAGGAAATGATTTTTTTTAAACC
ACCAATGCTGCAGGAAAAAGCAACATATTTAAGTTATCCAATAACACCTATCCAATAATTGTAAATCATTATCAT
GACATGGTAGAGTTGTTTATATTTCTTTTCTTTTAGGTGAAACACCATTCAAAGTCGTAGTCAAATCTCTTTCA
CCTAAAGAGTTGGTCCGGATACATGTCCCTAAACCTTTGGACAGGAATGATGGAACATTTTGTATGAGATATAGG
ATGTATGAAACTGTGATGAAGGCTGAAGATAGAGGTCCTTTATGGTGATGAACATGTGGCTCAGTCTCCCTAT
ATTTTGAAGGACCACTGTACCATGAGTACTGTAGTGTCCGGAAGATCCTCAGGCTGGCAGAAGACTCTTTCT
TGTCCAACCAAGGAACACAGATTGCAAAAGATTTTGCTTCCCTTTCCAGCATCAATCTCCAGCAAATGCTAAAA
GAAGTCCCCAAAAGGTTTGGGGATGAGAGAGGTGCCATTGTTTATTACAGATTCTCAATAACCATGTTTACCGG
AGATCTTTAGGGAAATACACAGACTTCAAGATGTTCTCTGATGAGATTTTGTATCATTGACAAGAAAGGTCCTT
CTCCAGATTTAGAATTTTATGTTAATCTTGGAGATTGGCCCTTGGAGCATCGAAAAGTCAATGGAACCCCTAGC
CCCATACCTATCATTTTATGGTGTGGCTCTCTGGATTCAAGAGATGTTGTCTTCCAACGTATGACATCACCAC
TCCATGCTTGAAGCCATGCGGGGTGTTACAAATGATCTCTCTCTATTTCAGGGAATACAGGGCCTTCTCTGATC
AATAAACAGAGAGAGCTTTCTTCAAGGTAGAGACAGCCGAGAGGAGAGGCTCCAGTTGGTACAGCTGTCCAAA
GAAAATCCTCAGCTACTAGATGCAGGAATTACAGGATATTTCTTTTCCAAGAGAAAGAAAAGGAGCTTGGAAAA
GCCAAGTTGATGGGTTTCTTTGATTTCTTTAAGTACAAGTATCAAGTAAATGTGGATGGGACCGTGGCTGCTTAC
AGATATCCATATCTCATGCTGGGCGACAGTCTGGTTTAAAGCAGGACTCGCCATATTATGAACATTTCTACATG
GCAC TAGAACCTTGGGAAGCATTTATGTTCCAATTAAGAAATCTGAGTGATTATTTAGAGAAAGTTAAATGGGCT
AAGGAAATGATGAAGAAGCCAAGAAGATTGCAAAAGAAGGACAGTTGATGGCTAGGGACCTACTACAGCCACAC
AGGCTTTACTGCTACTATTACCAAGTACTGCAGAAATATGCCGAGCGCCAGTCCAGCAAACCCGAAGTACCTGAT
GGAATGGAACCTTGTTCTCAGCCAGAAGATAGCACAGCCATCTGCCAGTGCCACAGGAAAAAGCCTTCAAGAGAA
GAACCTTGAAGTCAAGCCAGAATCACACTCCTGTGTATCCCGGTACACTTTAAGGAAAGATTGAATCTAAGCTGT
GAAGGACAGTATAGAAGACTGCACCAAGTGGACTAGTTCTCCCGGTGGCTTTATATATGTAGATGGATATAGCAG
TACTGGTTGAGTATCCCTCATCTGAAATGCTTAGGACCAGGAGTGTTCAGGCTTCAGATTTTTTAAGATTTGGG
AATATTTGCATGTACATAATGAGGTATCTTGGGGATGAGATCCAACTCTAAACACAAAATTCAATTTATATTTTAT
ATATACCTTGTTTACATACCTGAAGGTAATTTTATATAATATTTTAAATAATTTGTGCATGAAACAAAGTTTGT
ATACATTGAAGTGTGAGAAAGCAAGGTGCTACTATCTTAGCGACCCAAGTGGTGGTGTGAGCACTCAAAAAGTT
TTGGATTTTGGGGTATTTTCAATTTTATGTTTGTATGAGGAATGTTCAACCTGTATTTGAACAAGCATTTACCA
AATATCATTGAATATTAATATCTTTTGGTAAAACTGCTATTATCAGCATCATAGTTTCTCTAAAAAGAAAATCT
TGGGGATCATAGCCGATAGAGAGACTTGCTAAAATATAAATCAGCCTCTGCAAACTGTTTACATATTTATTGGT
TTACATATTTTATTGGTTTATTTCTATCCCTGTTCACTTTTCTCTTCCACTTCCAATATGAAGAGAAAATAT
TTGTTCAAGGTTGTCCCCCGCCCCCGTCACTGCATAATTTCTCTCTTACAAGCTGCTTTTGGCTTTTCAATTA
TAACAGCTTCTTTTGTAGAGGTCTGATAAGGATATTTAAGGAAGAGAGAAAGTACTGTTTATTAAGGTGGCAT
GGAGACTGTGGAGGGAATATTTTAAAGCACTACTCATATCCTTTAAACTAAATTTTGCCAAAGCCCGAGACAA
CATTAAGGAGAAATTTGATCCTTAAGTTAGTAATTTCAAATCTATCTGAGTTGTATACCCATCAAAGACAATACAG
TTATTAACATAGATGAAGGTATGCTATAGGCATCATTCATTATCTCTATATTGAATAGGTGAAAGATAACTGTAG
TCAGGTGAAAGGCTATCATCATTTTAAAGTGAAGGGGATCCTTGAAACACTGAAACCTCTACACAACTCT
TCAGGAAGCCTGCTATCTTGGGATCTACTAATAATAGGCCAAGAACAAAGGCAAGCATCCATCTCCATCCACC
ACTTTTCTATTTTCACTGGGTGTCTTGTACGATGAAGACTTTGGAATTTCTTTCTCTTTTAGGACAGGGTCA
GGATTTAGGACTCATAGCCTGAAAGCTCATTACATACTCCTTGAACCATCAGTCCAAGGTTTCACTTCACTAAAG
TGCATGTTCTAAAACAAGAGCTATCCTCATTCCAAATTTTAAATATGTACTCTGGCCGGTGTGAGTGGCTCACG
CCTGTAATCCCAGCACTTTTGGCAGGCCGAGATGGGCGGATCTTTTGGGTGAGGATTTGAGACCAAGCTGGCCA
ACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCAGGCATGGTGGCATTTGCCTGTAAATCCCAGCT
ACTCGGGAGGCTGAGGCAGGAGAACTCACTGAACCTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTACACCACTG
CACTCCAGCCTGGGTGACAGAGTGAGACTCCATCTCAAAACTGAAAATAAAAAATAAAAATATGTATTCTCCTAA
CTGAAATATTTACTTAATCTGGAACCAATGTAACATTTTAAAGTGGTTACATCTATTCTTGCTGAAGAACAA
TAAACAGAATTTTGTGACTAAGCATAACCAATTTTCAAGACAGTCTAATCAATGCCAAGTATCCAGGCAAACTC
TAATACCCATCCATTGTGCAAAACCAAGCAGCAAGTATTAATAAGAGCAAGCTGTCTGAGCCCATACCTA
ATGAATTTGTGCTTAAATATTGTACATTGTGTTTGGGCTGTGCAAACTGGGATTATGGCAAGAAAGGTTGCC
TAACCTCATACCTTTCTGCCTCAAATTTCCAGGTGCTAAAGGCTAATGGCATTTTAAACATCTTACATTTTAAAAA
TTTATATTGCCTCTGCCAAACAGGCCTAATAGTTAAAGCAAGTTGAGACAAACAGGCAGATTCACTGTGTGGA
ACAGGAAGGATGTGCTTTAAAAAAGGTGGAATCCCTCAAAAAATTTCTATAGGGAGACAGCAGCCTTAATCTACA
TAATTTCTCATCTGCCAATTCAGCCGACGCTTTAAGAGTTAGTGTTAATGGCTTTCTGGTTTGAAGCAAAA
ATGCATCTATGTGGTTGAAAGTTTGGGAGGAGATTACCAATATCTGAGGAGAAGATGGAGTGAAGGGAATCTT
ACTTTTGTCTTTTATACCTTTCTATAATATTTAGATTTTTTTTACTGTAAGTATGGATCAAATTGCAAAATAAG
AAAAATGCCAACCTTAGAAAAGACAATAAATGCACAAAAGATATAAACAGGAACAGCAAATATTTATTTTTTC
CATTTTGTCTTTTTTAAATCTATGTTTAGAATTTATATCTTGGGACTTATGTATATATACCTTTTAAATAAA
ATAAATTTTCTAAATAAAAGTTG

126/550

FIGURE 126

MVELFIFLFLGGETPFKVVVKSLSPKELVRIHVPKPLDRNDGTFLMRYRMYETVDEGLKIEVL
YGDEHVAQSPYILKGPVYHEYCECPEDPQAWQKTLSCPTKEPQIAKDFASFPSINLQQMLKEV
PKRFGDERGAIVHYTILNNHVYRRSLGKYTDFKMFSDEILLSLTRKVLLPDLEFYVNLGDWPL
EHRKVNGTPSPIPIISWCGSLDSRDVVLPTYDITHSMLEAMRGVTNDLLSIQGNTGPSWINKT
ERAFFRGRDSREERLQLVQLSKENPQLLDAGITGYFFFQEKELGKAKLMGFFDFFKYKYQV
NVDGTVAAYRYPYLMGLGDSLVLKQDSPYYEHFYMALEPWKHYVPIKRNLSDLLEKVKWAKEND
EEAKKIAKEGQLMARDLLQPHRLYCYYYQVLQKYAERQSSKPEVRDGMELVPQPEDSTAICQC
HRKKPSREEL

Important features of the protein:**Signal peptide:**

amino acids 1-16

N-glycosylation sites.

amino acids 250-254, 363-367

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 444-448

N-myristoylation site.

amino acids 208-214, 319-325, 388-394

Endoplasmic reticulum targeting sequence.

amino acids 448-453

Mitochondrial energy transfer proteins signature.

amino acids 25-34

127/550

FIGURE 127

AGCCGTCGGAGGGAGCCGGAGCGCTTCTCCCGAGTTGGTGATAGATTGGTGGTCATCCAACAT
GCAGAAATGAATGAGCAGTGAAAAGCAGCAGAGCCGATGGGTCATGAGGATGTAAGTGCGTTT
GAAGGCTTCCACACCCCTCTACTCCAGGAATCATGAATAAACTGGAGGATAAGCAGGACCAGAT
GATACCATGAAAGAGAAGTTTACAGGCCCTCTATTGCCAACTGTAAAGTTTCTGCTGATCTTG
GCACTGACCGAAGCGCTGGCATTGTCATCCAGGAACCATCTCCCAGGGAATCTCTTCAGGTC
CTCCCTTCAGGCACTCCCCCGGAACCATGGTGACAGCACCCACAGCTCTACCAGACATACT
TCTGTGGTGATGCTGACCCCCAATCCCGATGGACCCCCCTCACAGGCTGCAGCTCCCATGGCA
ACACTGACACCCCGTGACAGAGGGGCACCCTCCTACGCACACCATCTCCACCATCGCTGCGACA
GTAACCGCCCCCTATTCTGAAAGCTCCCTGTCCACAGGGCCCGCTCCAGCAGCCATGGCAACC
ACATCCTCCAAGCCAGAGGGCCGCCCCGAGGGCAGGCTGCCCCCACCATCCTGCTGACAAAG
CCACCGGGGGCCACCAGCCGCCCCACCACAGCGCCCCCGCACTACCACACGCAGGCCCCCC
AGGCCCCCAGGCTCTTCCCGAAAAGGGGCTGGTAATTCATCACGCCCTGTCCCGCCTGCACCT
GGTGGCCACTCCAGGAGTAAAGAAGGACAGCGAGGACGAAATCCAAGCTCCACACCTCTGGGG
CAGAAGCGGGCCCTGGGGAAAATCTTTCAGATCTACAAGGGCAACTTCACAGGGTCTGTGGAA
CCAGAGCCCTCTACCCTACCCCCAGGACCCCACTCTGGGGCTACTCCTCTTACCACAGCCC
CAGACAGTGGCTGCGACCACAGTGCCAGCAATACCTCATGGGCACCCACCACCACCTCCCTG
GGGCTGCAAAGGACAAGCCAGGCCTTCGCAGAGCAGCCCAGGGGGGTGGTTCTACCTTACC
AGCAAGGAGGACACCAGATGCCACAGCAGCCTCAGGTGCCCTGTCACTCCACAAGCTGCC
CCAGTGCTTCTCAGCGCCCCACCACGGTGACCCACAGGATGGCCCCAGCCATAGTGACTCT
TGGCTTACTGTTACCCCTGGCACCCAGCAGACCTCTGTCTACCAGCTCTGGGGTCTTCACGGCT
GCCACGGGGGCCACCCCAGCTGCCTTCGATACCAGTGTCTCAGCCCCCTTCCAGGGGATTCT
CAGGGAGCATCCACAACCCCAAGCTCCAACCCATCCCTCCAGGGTCTCAGAAAGCACTATT
TCTGGAGCCAAGGAGGAGACTGTGGCCACCCTCACCATGACCGACCGGGTGCCAGTCCTCTC
TCCACAGTGGTATCCACAGCCACAGGCAATTTCTCAACCGCCTGGTCCCCGCCGGGACCTGG
AAGCCTGGGACAGCAGGGAACATCTCCCATGTGGCCGAGGGGGACAAACCGCAGCACAGAGCC
ACCATCTGCCTGAGCAAGATGGATATCGCCTGGGTGATCCTGGCCATCAGCGTGCCCATCTCC
TCCTGCTCTGTCTGCTGACGGTGTGCTGCATGAAGAGGAAGAAGAAGACCGCCAACCCGGAG
AACAACCTGAGCTACTGGAACAACACCATCACCATGGACTACTTCAACAGGCATGCTGTGGAG
CTGCCCAGGGAGATCCAGTCCCTTGAAACCTCTGAGGACCAGCTCTCAGAGCCCCGCTCCCCA
GCCAATGGCGACTATAGAGACACTGGGATGGTCCTTGTTAACCCCTTCTGTCAAGAAACACTG
TTTGTGGGAAACGATCAAGTATCTGAGATCTAACTACAGCAGGCATCACTTTGCCATTCCGTA
TTTTTCGTCTCTAAATTATAAATATACAAATATATATATTATAAATATAACCTTGTGTAACCC
TGACTTAATGAGAAACATTTTCAGCTTTTTTTTCTATGAATTGTCAACATCTTTTTTTACAAGT
GTGGTTTAAAAAAAAAAAAAACTTTACAGAATGATCTGTGGCTTTATAAAATAAAGGTATTTCT
AAGCAAAAAAAAAAAAAAAAAA

128/550

FIGURE 128

MKRSLQALYQCQLLSFLLILALTEALAFAIQEPSPRESLQVLPSGTPPGTMVTAPHSSTRHTSV
VMLTPNPDGPPSQAAAPMATLTPRAEGHPPTHISTIAATVTAPYSESSLSTGPAPAAMATTS
SKPEGRPRGQAAPTILLTKPPGATSRPTTAPPRTTTRRPPRPPGSSSRKGAGNSSRPVPPAPGG
HSRSKEGQRGRNPSSTPLGQKRPLGKIFQIYKGNFTGSVEPEPSTLTPRTPLWGYSSSPQPQT
VAATTVPNTSWAPTTTSLGPAKDKPGLRRAAQGGGSTFTSQGGTPDATAASGAPVSPQAAPV
PSQRPHHGDPQDGPSHSDSWLTVTPGTSRPLSTSSGVFTAATGPTPAAFDTSVSAPSQGIPQG
ASTTPQAPTHPSRVSESTISGAKEETVATLTMTRVPSPLSTVVSTATGNFLNRLVPAGTWKP
GTAGNISHVAEGDKPQHRATICLSKMDIAWVILAIISVPISSCSVLLTVCCMKRKKKTANPENN
LSYWNNTITMDYFNRHAVELPREIQSLETSEDQLSEPRSPANGDYRDTGMVLVNPFCQETLFV
GNDQVSEI

Important features of the protein:**Signal peptide:**

amino acids 1-28

Transmembrane domain:

amino acids 469-487

N-glycosylation sites.

amino acids 178-182, 223-227, 261-265, 446-450, 504-508, 509-513

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 495-499

N-myristoylation sites.amino acids 44-50, 48-54, 175-181, 222-228, 279-285, 286-292,
288-294, 296-302, 351-357, 374-380, 427-433, 442-448**TonB-dependent receptor proteins signature 1.**

amino acids 1-44

129/550

FIGURE 129

AGGCGAGGCGCGGCGCCGCTGCACACACGCACACGGAGCT**ATG**GGGTGCCATGTTGCCACCAG
CTGCCACGTGGCCTGGCTTTTGGTGCTGATCTCTGGATGCTGGGGCCAGGTGAACCGGCTGCC
CTTCTTACCAACCACTTCTTTGATACATACCTGCTGATCAGCGAGGACACGCCTGTGGGTTC
TTCTGTGACCCAGTTGCTGGCCCAAGACATGGACAATGACCCCTGGTGTGGCGTGTCTGG
GGAGGAGGCCTCTCGCTTCTTTGCAGTGGAGCCTGACACTGGCGTGGTGTGGCTCCGGCAGCC
ACTGGACAGAGAGACCAAGTCAGAGTTCACCGTGGAGTTCTCTGTCAGCGACCACCAGGGGGT
GATCACACGGAAGGTGAACATCCAGGTCGGGGATGTGAATGACAACGCGCCACATTTACAA
TCAGCCCTACAGCGTCCGCATCCCTGAGAATACACCAGTGGGGACGCCCATCTTCATCGTGAA
TGCCACAGACCCCGACTTGGGGGCAGGGGGCAGCGTCCTCTACTCCTTCCAGCCCCCTCCCA
ATTCTTCGCCATTGACAGCGCCCGCGGTATCGTCACAGTGATCCGGGAGCTGGACTACGAGAC
CACACAGCCCTACCAGCTCACGGTCAACGCCACAGATCAAGACAAGACCAGGCCTCTGTCCAC
CCTGGCCAACCTTGGCCATCATCATCAGATGTCCAGGACATGGACCCATCTTCATCAACCT
GCCTTACAGCACCAACATCTACGAGCATTCTCCTCCGGGCACGACGGTGCGCATCACCGC
CATAGACCAGGATAAAGGACGTCCCCGGGGCATTGGCTACACCATCGTTTCAGGGAATACCAA
CAGCATCTTTGCCCTGGACTACATCAGCGGAGTGCTGACCTTGAATGGCCTGCTGGACCGGGA
GAACCCCTGTACAGCCATGGCTTCATCCTGACTGTGAAGGGCACGGAGCTGAACGATGACCG
CACCCCATCTGACGCTACAGTCACCACGACCTTCAATATCCTGGTTATTGACATCAATGACAA
TGCCCCGGAGTTCAACAGCTCCGAGTACAGCGTGGCCATCACTGAGCTGGCACAGGTCGGCTT
TGCCCTTCCACTCTTCATCCAGGTGGTGGACAAGGATGAGAATTTGGGCCTGAACAGCATGTT
TGAGGTGTACTTGGTGGGGAACAACCTCCCACTTTCATCATCTCCCGACCTCCGTCCAGGG
GAAGGCGGACATTGATTCGGGTGGCCATCCCACTGGACTACGAGACCGTGGACCGCTACGA
CTTTGATCTCTTTGCCAATGAGAGTGTGCCTGACCATGTGGGCTATGCCAAGGTGAAGATCAC
TCTCATCAATGAAAATGACAACCGGCCCATCTTCAGCCAGCCACTGTACAACATCAGCCTGTA
CGAGAACGTACCGTGGGGACCTCTGTGCTGACAGTCCTGGTGAGTCCCCGCTTCACTGCAGG
GCCACTGAGCTCTCCAGGGCCGACTGTGGTGAGGCACCCAGAGGGATTTTGTCCAAGGGACCT
CAGCAATCAGGGAAGGAGGCACCCCCAAATCCCTGAGCTGTGTTTGTGGTGTAT**TAA**AATAAA
GTTTTTGGACTCTTCAGGAAGGGGCTCCCTTGACCTAGGTTGCAATATGGAAAAGGAGCCAAC
CTGAGGGGTGACGAGACTGAGCTGAGGACACTGGTTTTCTGCCTTTCCCTGAGAGAGACTCAG
TGAGGGTGGGCTGGGAGCCCTGGAAGCCCCCTCAAATGGGTGGGAAGGTGCCAGCCATCCTTG
AGAAGGGCAACCTCTCCATGTGAGCACAGGCACCAGAGAGGGGCAGGCGCCTGGAGGGTACC
GGGGCACCCCACTGCCCATGGCTGGACTTGCCCTTTGACAAGGGGCCCTCCAGTGTCATT
TGTATCTGTCAGTACTCTTGGTTGCAAGGGACAGAAACCCTTAAGTAGTTCAAGCAAAAAAGG
ATTGGCTCATGTAACCAAAAGTATAAGTGATTTAGGCCGGGCTCGGTGGCTCACGCCTGTC
ATCCAACACCTTGAGAAAGCCGAGGTGGGCGGATCACTTGAGGTCGGGAGTTTGAGACCAGCC
TGGCCAACATGGCAAAACCCCGTCTCTACTAAAAATACAAAATTAGCCGGGTGTGGTGAC
ACGCCTGTAGTCCCAGCTACTAGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCAGGAGGCGG
AGGTTGCAGTGAGCCGAGATTGTGTCACTGCCCTCCAGCCTGGGCGACAGAGCCAGATTCTGT
CTC

130/550

FIGURE 130

MGCHVATSCHVAWLLVLISGCWGQVNRLPFFTNNHFFDTYLLISEDTPVGSSVTQLLAQDMDND
PLVFGVSGEEASRFFAVEPDTGVVWLRQPLDRETKSEFTVEFSVSDHQGVITRKVNIQVGDVN
DNAPT FHNQPYSVRIPENTPVGTPIFIVNATDPDLGAGGSVLYSFQPPSQFFAIDSARGIVTV
IRELDYETTQAYQLTVNATDQDKTRPLSTLANLAIITDVQDMDPIFINLPYSTNIYHSPPG
TTVRIITAIDQDKGRPRGIGYTIVSGNTNSIFALDYISGVLTNLGLLDRENPLYSHGFILTVK
GTELNDDRTPSDATVTTTTFNILVIDINDNAPEFNSSEYSVAITELAQVG FALPLFIQVVDKDE
NLGLNSMFEVYLVGNNSHHFIISPTSVQ GKADIRIRVAIPLDYETVDRYDFDLFANESVPDHV
GYAKVKITLINENDNRPIFSQPLYNISLYENVTVGTSVLTVLVSPRFTAGPLSSPGPTVVRHP
EGFCPRDLSNQGRRHPQIPELCLLVY

Important features of the protein:**Signal peptide:**

amino acids 1-23

Transmembrane domain:

amino acids 355-374

N-glycosylation sites.amino acids 155-159, 206-210, 349-353, 393-397, 434-438, 466-470,
472-476**N-myristoylation sites.**

amino acids 2-8, 49-55, 162-168, 270-276, 278-284, 316-322

Amidation site.

amino acids 515-519

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 11-22

Leucine zipper pattern.

amino acids 298-320

PTS HPR component serine phosphorylation site signature.

amino acids 377-393

Cadherins extracellular repeated domain signature.

amino acids 120-131, 336-347

Cadherins extracellular

amino acids 120-144, 336-360

131/550

FIGURE 131

GTGGGCCGCCCTGCTGCTGCCGTCCATGCTGATGTTTGCGGTGATCGTGGCCTCCAGCGGGC
TGCTGCTCATGATCGAGCGGGGCATCCTGGCCGAGATGAAGCCCCTGCCCTGCACCCGCCCCG
GCCGCGAGGGCACAGCCTGGCGCGGGAAAGCCCCCAAGCCTGGGGGCCTGTCCCTCAGGGCTG
GGGACGCGGACTTGCAAGTGCGGCAGGACGTCCGGAACAGGACCCTGCGGGCGGTGTGCGGAC
AGCCAGGCATGCCCCGGGACCCCTGGGACTTGCCGGTGGGGCAGCGGCGCACCCCTGCTGCGCC
ACATCCTCGTAAGTGACCGTTACCGCTTCCTCTACTGCTACGTCCCCAAGGTGGCCTGCTCTA
ACTGGAAGCGGGTGATGAAGGTGCTGGCAGGCGTCCTGGACAGCGTGACGTCCGCTCAAGA
TGGACCACCGCAGTGACCTGGTGTTCCTGGCCGACCTGCGGCCTGAGGAGATTCGCTACCGCC
TGCAGCACTACTTTAAGTTCCTGTTTGTGCGGGAGCCCTTGGAACGCCTCCTCTGCTTACC
GCAACAAGTTTGGCGAGATCCGAGAGTACCAGCAACGCTATGGGGCTGAGATAGTGAGGCGGT
ACAGGGCTGGAGCGGGGCCAGCCCTGCAGGCGACGATGTCACATTCCCCGAGTTCCTGAGAT
ACCTGGTGGATGAGGACCCTGAGCGCATGAATGAGCATTGGATGCCCCGTGACCACCTGTGCC
AGCCTTGTGCCGTGCACTATGACTTTGTGGGCTCCTATGAGAGGCTGGAGGCTGATGCAAATC
AGGTGCTGGAGTGGGTACGGGCACCACCTCACGTCCGATTTCCAGCTCGCCAGGCCTGGTACC
GGCCAGCCAGCCCCGAAAGCCTGCATTACCACTTGTGCAGTGCCCCCGGGCCCTGCTGCAGG
ATGTGCTGCCTAAGTATATCCTGGACTTCTCCCTCTTTGCCTACCCACTGCCTAATGTCACCA
AGGAGGCGTGTCAGCAGTGAACCATGGGTGTGGGGCCAGCAGCTGGTGGGGACTGGTTTCAACG
CCAGCTTTCTGTGCTTCTGCCTGTCATTCGGAGAACTCTGGCTCTGGGGCTTGGGGCTTCTC
AGGATCCTGGATGGCAGAGACTGCCCTCAGAAGTTCCTTGTCAGGGTGGGCACCCACAGTGA
CTCAGAGGACAGGGCTAGGCAGGAGACCTGCTGCTCCTCATTGGGGGGATCTCTTGGGGGGCA
GACACCAGTTTGCCAATGAAGCAACACATCTGATCTAAAGACTGGCTCCAGACCCCGGGCTGC
CAGGATTATGCAGTCCACTTGGTCTACCTTAATTTAACCTGTGGCCAACTCAGAGATGGTAC
CAGCCAGGGGCAAGCATGACCAGAGCCAGGGACCCTGTGGCTCTGATCCCCCATTTATCCACC
CCATGTGCCTCAGGACTAGAGTGAGCAATCATACCTTATAAATGACTTTTGTGCCTTTCTGCT
CCAGTCTCAAAATTTCTACACCTGCCAGTTCCTTACATTTTCCAAGGAAAGGAAAACGGAA
GCAGGGTTCTTGCCTGGTAGCTCCAGGACCCAGCTCTGCAGGCACCCAAAGACCCTCTGTGCC
CAGCCTCTTCCTTGAGTTCTCGGAACCTCCTCCCTAATTCTCCCTTCCTTCCCCACAAGGCCT
TTGAGGTTGTGACTGTGGCTGGTATATCTGGCTGCCATTTTCTGATGCATTTATTTAAATTT
TGTAATTTTGTATAGAACCCTTGTAAGGGCTTGTTCCTAATAGCTGACTTTTAAATAAAG
CAGTTTTATATAT

132/550

FIGURE 132

MLMFAVIVASSGLLLMIERGILAEMKPLPLHPPGREGTAWRGKAPKPGGLSLRAGDADLQVRQ
DVRNRTLRAVCGQPGMPRDPWDLVPGQRRTLRHILVSDRYRFLYCYVPKVACSNWKRVMKVL
AGVLDSVDVRLKMDHRSDLVFLADLRPEEIRYRLQHYFKFLFVREPLERLLSAYRNKFGEIRE
YQORYGAEIVRRYRAGAGPSPAGDDVTFPEFLRYLVDEDPERMNEHWMPVYHLCQPCAVHYDF
VGSYERLEADANQVLEWVRAPPHVRFPARQAWYRPASPESLHYHLC SAPRALLQDVLPKYILD
FSLFAYPLPNVTKEACQQ

Important features of the protein:

Signal peptide:

amino acids 1-23

N-glycosylation sites.

amino acids 67-71, 325-329

Tyrosine kinase phosphorylation sites.

amino acids 152-159, 183-183

N-myristoylation sites.

amino acids 89-95, 128-134

133/550

FIGURE 133

CGGCAGTTCTGGCCCCTGCAGCTGGAGGTACCCTGAGTTCTGAGGGTCGTAAGTGCTGTTTCTG
GTATTCTCATCGCGGTCACTCTACCGGTGTGGACAAGTAAAGTTTGAATCAGCTTCTCCATG
GCCTGGGCACCAAGTTCCCGGCTGAGCCATTTTCCTTTTGGCTAAAAGTCCCCGCCAGAGGCC
AATTTCGTCGCGGCGGCGGTGGAGATCGCAGGTCGCTCAGGCTTGCAGATGGGTCAAGGGTTGT
GGAGAGTGGTCAGAAACCAGCAGCTGCAACAAGAAGGCTACAGTGAGCAAGGCTACCTCACCA
GAGAGCAGAGCAGGAGAATGGATGCGAGCAACATTTCTAACACCAATCATCGTAAACAAGTCC
AAGGAGGCATTGACATATATCATCTTTTGAAGGCAAGGAAATCGAAAGAACAGGAAGGATTCA
TTAATTTGGAAATGTTGCCTCCTGAGCTAAGCTTTACCATCTTGTCTACCTGAATGCAACTG
ACCTTTGCTTGGCTTCATGTGTTTGGCAGGACCTTGCGAATGATGAACTTCTCTGGCAAGGGT
TGTGCAAATCCACTTGGGGTCACTGTTCCATATACAATAAGAACCCACCTTTAGGATTTTCTT
TTAGAAAATTGTATATGCAGCTGGATGAAGGCAGCCTCACCTTTAATGCCAACCAGATGAGG
GAGTGAACACTTTTATGTCCAAGGTATCCTGGATGATTGCGCAAAGGAAATAGCAAAGTTTA
TCTTCTGTACAAGAACTAAATTGGAAAAAAGTGAATCTATCTTGATGAAAGGAGAGATG
TCTTGGATGACCTTGTAACATTGCATAATTTTAGAAATCAGTTCTTGCCAAATGCACTGAGAG
AATTTTTTTCGTCATATCCATGCCCCTGAAGAGCGTGGAGAGTATCTTGAACTCTTATAACAA
AGTTCTCACATAGATTCTGTGCTTGCAACCCTGATTTAATGCGAGAACTTGGCCTTAGTCCTG
ATGCTGTCTATGTACTGTGCTACTCTTTGATTCTACTTTCCATTGACCTCACTAGCCCTCATG
TGAAGAATAAAATGTCAAAAAGGGAATTTATTGAAATACCCGTCGCGCTGCTCAAAATATTA
GTGAAGATTTTGTAGGGCATCTTTATGACAATATCTACCTTATTGGCCATGTGGCTGCATTAAA
AAGCACAATTGCTAGGACTTCAGTTTTTACTTCAGACTAAAGCTACCCAAGGACTTAGCAGAT
ATGGGGGTTACATCAGTGCTGGTCATTGTAGCCTGAGTATACAATCAAGCTTCAGTGTGCAAC
CTTTTTTTCTTTTGCCATTTTCTATTTTAGTAATTTCTTTGGGGAATAATAATTTTGCAGA
ATTTTTCCTAATTTTGTATACGTTTTGCACAAAGCAGAGCCACTGTCTAACACAGCTGTT
AACGAATGATAAACTGACATTATACTCTAAAAGATGGTGTATTTGTGCATTAGATTTGCCTGA
AAAACCTTTATCCATTTCCATTCTTTATACAAATACCATGTAATGTGTACATATTTAACTAAAG
AGATTTATAGTCATAATTATTTTATTGTAAAGATTTTAACTAAAGTTTTTCCTTTTCTCTC

134/550

FIGURE 134

MGQGLWRVVRNQQLQQEGYSEQGYLTREQSRRMDASNISNTNHRKQVQGGIDIIYHLLKARKSK
EQEGFINLEMLPPELSFTILSYLNATDLCLASCVWQDLANDELLWQGLCKSTWGHCSIYNKNP
PLGFSFRKLYMQLDEGSLTFNANPDEGVNYFMSKGILDDSPKEIAKFIFCTR TLNWKKLRIYL
DERRDVLDDLVTLHNFRNQFLPNALREFFRHHAPEERGEYLETLITKFSHRFCACNPDL MRE
LGLSPDAVYVLCYSLILLSIDLTSPHVKNKMSKREFIRNTRRAAQNISEDFVGHLYDNIY LIG
HVAA

Important features of the protein:

Transmembrane domain:

amino acids 253-272

N-glycosylation sites.

amino acids 37-41, 87-91, 298-302

N-myristoylation site.

amino acids 110-116

135/550

FIGURE 135

GGCACGAGGGAGCCTCCGTTAGGGGGTGGGAAAGGACTTTGCCATAGGTCGCTGAGGCCACCA
TCTGCTCTCTTACTGGCCAAGGGCGTAAAAAGATAGTCTTCCCATTAGCTAGAGAGCAAACCC
CAGAAAGCCTATTGGCTGCGCCGTCCGCGGGCCTTGGTCCGCTTTGAAGGCGGGCTGCGGCTG
CGAGAGGAGGGCGGGCGGGAGGCTAGCTGTTGTCGTGGTTGCTCGGAGGCACGTGTGCAGTCC
CGGAAGCGGCGAGGGGAACTGCTCCGCGCGCGCCGCGGGAGGAGGAACCGCCCGGTCCTTTA
GGGTCCGGGGCCCGGCCGGGCCATGGATTCAATGCCTGAGCCCCGCTCCCGCTGTCTTCTGCTT
CTTCCCTTGCTGCTGCTGCTGCTGCTGCTGCTGCCGGCCCCGGAGCTGGGCCCCGAGCCAGGCC
GGAGCTGAGGAGAACGACTGGGTTTCGCTGCCAGCAAATGCGAAGTGTGTAAATATGTTGCT
GTGGAGCTGAAGTCAGCCTTTGAGGAAACCGGCAAGACCAAGGAGGTGATTGGCACGGGCTAT
GGCATCCTGGACCAGAAGGCCTCTGGAGTCAAATACACCAAGTCGGACTTGCGGTTAATCGAA
GTCAGTGAAGACCATTTGCAAGAGGCTCCTGGATTATAGCCTGCACAAGGAGAGGACCGGCAGC
AATCGATTTGCCAAGGGCATGTCAGAGACCTTTGAGACATTACACAACCTGGTACACAAGGG
GTCAAGGTGGTGTATGGACATCCCCTATGAGCTGTGGAACGAGACTTCTGCAGAGGTGGCTGAC
CTCAAGAAGCAGTGTGATGTGCTGGTGAAGAGTTTGAGGAGGTGATCGAGGACTGGTACAGG
AACCACCAGGAGGAAGACCTGACTGAATTCCTCTGCGCCAACCACGTGCTGAAGGGAAAAGAC
ACCAGTTGCCTGGCAGAGCAGTGGTCCGGCAAGAAGGGAGACACAGCTGCCCTGGGAGGGAAG
AAGTCCAAGAAGAAGAGCAGCAGGGCCAAGGCAGCAGGCGGCAGGAGTAGCAGCAGCAAACAA
AGGAAGGAGCTGGGTGGCCTTGAGGGAGACCCCAGCCCCGAGGAGGATGAGGGCATCCAGAAG
GCATCCCCTCTCACACACAGCCCCCTGATGAGCTCTGAGCCCCACCCAGCATCCTCTGTCTTG
AGACCCCTGATTTTGAAGCTGAGGAGTCAGGGGCATGGCTCTGGCAGGCCGGGATGGCCCCGC
AGCCTTCAGCCCCCTCTTGCCCTTGGCTGTGCCCTCTTCTGCCAAGGAAAGACACAAGCCCCAG
GAAGAACTCAGAGCCGTCATGGGTAGCCCACGCCGTCTTTCCCCCTCCCCAAGTGTTTCTCTC
CTGACCCAGGGTTCAGGCAGGCCTTGTGGTTTCAGGACTGCAAGGACTCCAGTGTGAAGTCAAG
GAGGGGCAGGTGTCAGAACTGGGCACCAGGACTGGAGCCCCCTCCGGAGACCAAACCTACCAT
CCCTCAGTCCTCCCCAACAGGGTACTAGGACTGCAGCCCCCTGTAGCTCCTCTCTGCTTACCC
CTCCTGTGGACACCTTGCACTCTGCCTGGCCCTTCCCAGAGCCCAAAGAGTAAAAATGTTCTG
GTTCTGATTTCTGAAAAAAAAAAAAAAAAAAATTCCT

136/550

FIGURE 136

MDSMPEPASRCLLLLLPLLLLLLLLLLPAPELGPSQAGAEENDWVRLPSKCEVCKYVAVELKSAF
EETGKTKEVIGTGYGILDQKASGVKYTKSDLRLIEVTETICKRLLDYSLHKERTGSGNRFAKGM
SETFETLHNLVHKGVKVMDIPYELWNETSAEVADLKKQCDVLVEEFEEVIEDWYRNHQEEDL
TEFLCANHVLKGKDTSCLAEQWSGKKGDTAALGGKKSKKKSSRAKAAGGRSSSSKQRKELGGL
EGDPSPEEDEGIQKASPLTHSPPEL

Important features of the protein:

Signal peptide:

amino acids 1-26

N-glycosylation site.

amino acids 153-157

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 227-231, 228-232

Tyrosine kinase phosphorylation site.

amino acids 142-150

N-myristoylation sites.

amino acids 36-42, 74-80, 86-92, 125-131, 222-228, 237-243,
250-256, 263-269

Amidation sites.

amino acids 212-216, 222-226

ATP/GTP-binding site motif A (P-loop).

amino acids 62-70

137/550

FIGURE 137

CACGCCTCCCGCTGCCAGCCCGGCACCGGGATCTTAATCAGTCACTATGAAAACCTCATTAGCT
CCACAGCAATGAGTCCCTCCACTGCTGAAGCTTGGCGCTGTGCTTAGTACCATGGCAATGATCT
CAAACCTGGATGTCCCAAACCTCTCCCATCCTTGGTGGGACTGAACACCACGAGGCTGTGCACTC
CGGATACCTTAACTCAGATTAGTCCTAAAGAAGGGTGGCAGGTGTACAGCTCAGCTCAGGATC
CTGATGGGCGGTGCATTTGCACAGTTGTTGCTCCAGAACAAAACCTGTGTTCCCGGGATGCCA
AAAGCAGGCAACTTCGCCAACTACTGGAAGAGGTTTCAGAACATGTCCAGTCTATTGAAGTCT
TAAACTTGAGAACTCAGAGAGATTTCCAATATGTTTTAAAAATGGAAACCCAAATGAAAGGGC
TGAAGGCAAAATTTTCGGCAGATTGAAGATGATCGAAAGACACTTATGACCAAGCATTTTTCAGG
AGTTGAAAGAGAAAATGGACGAGCTCCTGCCTTTGATCCCCGTGCTGGAACAGTACAAAACAG
ATGCTAAGTTAATCACCCAGTTCAAGGAGGAAATAAGGAATCTGTCTGCTGTCTCACTGGTA
TTCAGGAGGAAATTTGGTGCCTATGACTACGAGGAACTACACCAAAGAGTGCTGAGCTTGGAAA
CAAGACTTCGTGACTGCATGAAAAAGCTAACATGTGGCAAACCTGATGAAAATCACAGGCCCGAG
TTACAGTCAAGACATCTGGAACCCGATTTGGTGTCTGGATGACAGACCCTTTAGCATCTGAGA
AAAACAACAGAGTCTGGTACATGGACAGTTATACTAACAATAAAATTTGTTTCGTGAATACAAAT
CAATTGCAGACTTTGTCTAGTGGGGCTGAATCAAGGACATACAACCTTCCTTTCAAGTGGGCAG
GAACTAACCATGTTGTCTACAATGGCTCACTCTATTTTAACAAGTATCAGAGTAATATCATCA
TCAAATACAGCTTTGATATGGGGAGAGTGCTTGCCCAACGAAGCCTGGAGTATGCTGGTTTTTC
ATAATGTTTTACCCTACACATGGGGTGGATTCTCTGACATCGACCTAATGGGCTGATGAAATCG
GGCTGTGGGCTGTGTATGCAACTAACCAGAATGCAGGCAATATTGTCTATCAGCCAACCTTAACC
AAGATACCTTGGAGGTGATGAAGAGCTGGAGCACTGGCTACCCCAAGAGAAGTGCAGGGGAAT
CTTTCATGATCTGTGGGACACTGTATGTCACCAACTCCCACTTAACTGGAGCCAAGGTGTATT
ATTCCATTTCCACCAAAACCTCCACATATGAGTACACAGACATTCCCTTCCATAACCAATACT
TTCACATATCCATGCTTGACTACAATGCAAGAGATCGAGCTCTCTATGCCTGGAACAATGGCC
ACCAGGTGCTGTTCAATGTCAACCTTTTCCATATCATCAAGACAGAGGATGACACATAGGCAA
ATGTGACATGTTTTCATTTGATTTAAACAGTGTGATTTGTGATAAACTCTATAAGACCCCTTCC
GTTTTTTTCTTCACTATTATTTTTTCATCATTTCTCCAAAGCAAAGCATTTTTATTGTAAAGTT
GGTGTTCAAAACATAGCTGAGCTTGTCTAACTTACCATGTTGGAAACACATCTTAACTTCT
AAATTTACAAGGCCTATCATGTCTTGTGATGAAAAGCACTAAAAAAAAAAAAAGAGTTTAAGT
GGCTAAAGTCATAGTTTTGCAAGAGATTAATGATCTGCCTTATATTAGAGTCAGAGACTAATG
GTGGCTTAAATGCACGAATGTCTTTTTTTTTTAAACTGTCATTTTTTTACTGTCTTTTGCTCCA
TCTCAGGAAATATTTTGGTAGGAATTAGGAGAACAAAAAGCACTTTTATCCCATTTATTTCTT
TAAAAAATGTAAGGATTTCAATTTATATTGAAAAATAATATTAATCATTTTGCTGTTAACACAA
TTCTCTGATGCGGTGCTGTACAGTCATTTTTTAAATCTCTTGCTAACATTTTATTGGCAGTATG
TATTTCTACCATTGTAACCACCATTTGTGCTATTGTATCTCTTCACTTCTGTGAAAGTAATATT
TTTTATAAANACACTGNAATTTTAAAAAAAAAAAAAAAAAAAAA

138/550

FIGURE 138

MSPPLLKLGAVLSTMAMISNWMSQTLPSLVGLNTTRLSTPDTLTQISPKEGWQVYSSAQDPDG
RCICTVVAPEQNLCSRDAKSRQLRQLLEKVQNMSQSIEVLNLRRTQRDFQYVLKMETQMKGKLA
KFRQIEDDRKTLMTKHFQELKEKMDELLPLIPVLEQYKTDAKLITQFKEEIRNLSAVLTGIQE
EIGAYDYEELHQRVLSLETRLRDCMKKLTGKLMKITGPVTVKTSGTRFGAWMTDPLASEKNN
RVWYMSYTNKIVREYKSIADFVSGAESRTYNLPFKWAGTNHVVYNGSLYFNKYQSNIIKY
SFDMGRVLAQRSLEYAGFHNVPYPTWGGFSDIDLMADEIGLWAVYATNQNAGNIVISQLNQDT
LEVMSWSTGYPKRSAGESFMICGTLYVTNSHLTGAKVYYSYSTKTSTYEYTDIPFHNQYFHI
SMLDYNARDRALYAWNNGHQVLFNVTLFHIKTEDDT

Important features of the protein:**Signal peptide:**

amino acids 1-16

N-glycosylation sites.

amino acids 33-37, 95-99, 179-183, 299-303, 465-469

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 215-219

Tyrosine kinase phosphorylation site.

amino acids 106-114

N-myristoylation sites.

amino acids 9-15, 31-37, 235-241, 239-245

139/550

FIGURE 139

GAAGCAGTGCAGAGAGGAGAGCGGAGCGGAGCTGCCGCTGAGCAAAGGCCTTCACCAATGGCCG
AGTCCCCCGGCTGCTGCTCCGTCTGGGCCCCGCTGCCTCCACTGCCTGTATAGCTGCCACTGGA
GGAAATGCCCCAGAGAGAGGATGCAAACCAGCAAGTGCGACTGTATCTGGTTTGGCCTGCTCT
TCCTCACCTTCCTCCTTTCCCTGAGCTGGCTGTACATCGGGCTCGTCCTTCTCAATGACCTGC
ACAACTTCAATGAATTCTCTTCCGCCGCTGGGGACACTGGATGGACTGGTCCCTGGCATTCC
TGCTGGTCATCTCTCTACTGGTCACATATGCATCCTTGCTATTGGTCCTGGCCCTGCTCCTGC
GGCTTTGTAGACAGCCCCTGCATCTGCACAGCCTCCACAAGGTGCTGCTGCTCCTCATTATGC
TGCTTGTGGCGGCTGGCCTTGTGGGACTGGACATCCAATGGCAGCAGGAGTGGCATAGCTTGC
GTGTGTCACTGCAGGCCACAGCCCCATTCTTTCATATTGGAGCAGCCGCTGGAATTGCCCTCC
TGGCCTGGCCTGTGGCTGATACCTTCTACCGTATCCACCGAAGAGGTCCCAAGATTCTGCTAC
TGCTCCTATTTTTTGGAGTTGTCCTGGTCATCTACTTGGCCCCCCTATGCATCTCCTCACCT
GCATCATGGAACCCAGAGACTTACCACCCAAGCCTGGGCTGGTGGGACACCGAGGGGGCCCCCA
TGCTGGCTCCCGAGAACACCCTGATGTCTTGCAGGAAGACAGCTGAATGCGGAGCTACTGTGT
TTGAGACTGATGTGATGGTCAGCTCCGATGGGGTCCCCCTTCCTCATGCATGATGAGCACCTCA
GCAGGACCACGAATGTAGCCTCTGTATTCCCAACCCGAATCACAGCCCACAGCAGTGACTTCT
CCTGGACTGAACTGAAGAGACTCAATGCTGGATCCTGGTTCTTAGAGAGGCGACCCTTCTGGG
GGGCCAAACCGCTGGCAGGCCCTGATCAGAAAGAGGCTGAGAGTCAGACGGTACCAGCATTAG
AAGAGCTATTGGAGGAAGCTGCAGCCCTCAACCTTTCCATCATGTTGACTTGCGCCGACCCC
CACAGAACCACACATACTATGACACTTTTGTGATCCAGACATTGGAGACTGTGCTGAATGCAA
GGGTGCCCCAAGCCATGGTCTTTTGGCTACCAGATGAAGATCGGGCTAATGTCCAACGACGGG
CACCTGGAATGCGCCAGATATATGGACGTCAGGGAGGCAACAGAACGGAGAGGCCCCAGTTTC
TTAACCTCCCCTATCAAGATCTGCCACTATTGGATATCAAGGCATTGCATAAGGATAATGTCT
CGGTGAACCTATTTGTAGTGAACAAGCCCTGGCTCTTCTCTCTGCTTTGGTGTGCAGGGGTGG
ATTCGGTCACCACCAACGACTGCCAGCTGCTGCAGCAGATGCGTTACCCTATCTGGCTTATTA
CCCCTCAAACCTACCTAATCATATGGGTCATTACCAATTGTGTTTCCACCATGCTGCTTTTGT
GGACCTTCCTCCTCCAAAGGAGATTTGTTAAGAAGAGAGGGAAAACCTGGCTTAGAAACAGCAG
TGCTGCTGACAAGGATCAACAATTTTCATGATGGAGTGAATGCCCTGCCCTGCTTCCCCACCCA
AGCCAGTCTACATTGCCCAAACAGCAAGGGTTGGAGAGTGGCTTAAGTGGAATGCTTCAGGGG
TGGTGGGTGCAAGTGGGGGGAGCTTTGCCAACAGGAGGTTTTGAACCATGAGGGCCCTCTGC
CCAGGTGATGGGCATTCCCTAAGCTGCTATGGAATCTGCTCCCTTTGGGGTTTTGACCTGAGA
TGTTTGGGAAGAGAGTGAGTAATGAGAAGTTTCTCCTCAAATGAACTAGAACAGAGGAAGTA
AAAGGGAGATTGCTCGGA

140/550

FIGURE 140

MAESPGCCSVWARCLHCLYSCHWRKCPRERMQTSKDCIWFGLLFLTFLLSLSWLYIGLVLLN
DLHNFNEFLFRRWGHWMDSLAFLLVISLLVTYASLLLVLALLRLCRQPLHLHSLHKVLLLL
IMLLVAAGLVGLDIQWQQEWHSRLRVSLQATAPFLHIGAAAGIALAWPVADTFYRIHRRGPKI
LLLLLFFGVVLVIYLAPLCISSPCIMEPRDLPPKPGLVGHRGAPMLAPENTLMSLRKTAECGA
TVFETDVMVSSDGVPFMLHDEHLSRTTNVASVFPTRIAHSSDFSWTELKRLNAGSWFLERRP
FWGAKPLAGPDQKEAESQTPALEELLEEEAALNLSIMFDLRRPPQNHTYYDTFVIQTLETVL
NARVPQAMVFWLPDEDRAVQRRAPGMRQIYGRQGGRTERPQFLNLPYQDLPLLDIKALHKD
NVSVNLFVVNKPWLFSLWCAGVDSVTTNDCQLLQQMRYPIWLITPQTYLIIWVITNCVSTML
LLWTFLLQRRFVKKRGKTGLETAVLLTRINNFME

Important features of the protein:**Transmembrane domains:**

amino acids 38-60, 83-107, 122-138, 156-173, 189-210, 484-506

N-glycosylation sites.

amino acids 349-353, 362-366, 415-419, 442-446

N-myristoylation sites.

amino acids 163-169, 413-419, 523-529

Leucine zipper pattern.

amino acids 93-115, 109-131

Glutamine amidotransferases class-II active site.

amino acids 1-13

141/550

FIGURE 141

GCCGCCGCCGCCGGGCTGGAGCCGAGCGCAGCAGCCACCGCCGCCGCCGCCAGAAATTTGGGTTGAACCGGAGC
 TGCCGGGAGGAAACTTTTTTCTTTTTTCCCCCTCCCTCCCGGGAGGAGGAGGAGGAGGAGGGGAAGCTGCCG
 CCGCGCCCAAGGCTCGTGGGCTCGGGGTGGCGCGGCCCGCAGAAGGGGCGGGGGCCTCGCCCCGCGAGGGGAGG
 CGCGCCGCCGGGGCCCCGAGAGGGGGCGTGAGGACCGCGGGCTGCTGGTGCGGCGGGCGGGCGCGTGTGCCCCG
 CCGAGGGGAGGGCGCCCGCCCGCTCCGCGCCGGCTGCGAGGAGAGGCGCGCGCGCGCAGGAGGATGCTACTT
 GGTGGCGGGGACAGGGGGTTGGCCGGCTGCGGGCACCTCTGCTGTCTCGCTGCTGGGGCTGCTGCTGCTGGC
 GCGCTCCGGCACCCGGGCGCTGGTCTGCCTGCCCTGTGACGAGTCCAAGTGCAGGAGGCCAGGAAGTGCCCGGG
 GAGCATCGTGAGGGCGTCTGCGGCTGCTGCTACACGTGCCCCAGCCAGAGGAACGAGAGCTCGGGCGGCACCTT
 CGGGATTACGGAACCTGCGACCGGGGGCTGCGTGTGTCTATCCGCCCCCGCTCAATGGCGACTCCCTCACCG
 GTAGCAAGCGGGGCTTGGCAAGATGAGAACTGGAATGATGACCAACTGCTTGGTTTAAACCATGCAATGAAAA
 CCTTATTGCTGGCTGCAATATAATCAATGGGAAATGTGAATGTAAACACCTTGAACCTGCAGCAATCCCTTTGA
 GTTTCCAAGTCAGGATATGTGCCTTTCAGCTTTAAAGAGAATTGAAGAAGAGAAGCCAGATTGCTCCAAGGCCCG
 CTGTGAAGTCCAGTTCTCTCCACGTTGTCTGAAGATTCTGTTCTGATCGAGGGTTATGCTCCTCCTGGGGAGTG
 CTGCTCCCTTACCCAGCCGCTGCGTGTGCAACCCCGCAGGCTGTCTGCGCAAAAGTCTGCCAGCCGGGAAACCTGAA
 CATACTAGTGTCAAAGCCTCAGGGAAGCCGGGAGAGTGCTGACCTTATAGTGCAAAACGATTTTCGGCGT
 GGACTGCAGGACTGTGGAATGCCCTCCTGTTGACGAGACCGCGTGTCCCCGGACAGCTATGAAACTCAAGTCAG
 ACTAAGTGCAGATGTTTGTGTAAGTGTGCAACAAGATGCGAGTGTCTCTGCGCTTATGTGGTTTCCCCGTGTG
 TGAGGTGGGATCACTCCCCGCATAGTCTCTCGTGGCGATGGGACACCTGGAAGAGTCTGTGATGCTTTGAATG
 TGTTAATGATACAAAGCCAGCTGCGTATTTAAACAATGTGGAATATTATGATGGAGACATGTTTGAATGGACAA
 CTGTGCGTTCTGTGATGCCAAGGGGGCGTTGCCATCTGCTTCACTGCCAGTGTGGTGAGATAAAGTGCAGAG
 GTACTACGTGCCCGAAGGAGAGTGTGCCAGTGTGTGAAGATCCAGTGTATCCTTTTAATAATCCCGCTGGCTG
 CTATGCCAATGGCTGATCCTTGGCCAGGAGACCGGTGGCGGGAAGACGACTGCACATTCTGCCAGTGCCTCAA
 CGGTGACGCGCACTGCGTGTGCGACCGTCTGCGGACAGACCTGCACAAACCTGTGAAAGTGCTGGGGAGTGTG
 CCTGTGTGCGAAGAACCAACCATCATCACAGTGTACCTGCATGTGGGAGTTATCAAACTGCATCTGAG
 AGGGAAGGACTGCATTAATGGTTTCAAACGCGATCACAATGGTTGTGCGGACTGTGATGCATAAACACCGAGGA
 ACTATGTTTCAAGCGTAAACAAGGCTGCACCTTGAAGTGTCCCTTCGGTTTCTTACTGATGCCCAAACTGTGA
 GATCTGTGAGTGCCGCCCCAAGGCCCAAGAGTGCAGACCATAATCTGTGACAAGTATTGTCCACTTGGATTGCT
 GAAGAATAAGCACGGCTGTGACATCTGTGCTGTGAAGAAATGTCCAGAGCTCTCATGTCAGTAAAGTGTGCCCTT
 GGGTTTCCAGCAGGACAGTCAAGGCTGCTTATCTGCAAGTGCAGAGAGGCTCTGCTTCACTGGGCGACCCAT
 CCTGTGCGGCATTTGTCTCACCGTGGATGGTTCATCATATAAAATGAGGAGAGCTGGCACGATGGGTGCCGGGA
 ATGCTACTGTCTCAATGGACGGGAAATGTGTGCCCTGATCACCTGCCCGTGCCTGCCTGTGGCAACCCACCAT
 TCACCTTGGACAGTGTCTGCCCATCATGTGCAGATGACTTTGTGGTGCAAGGCCAGGCTCAGTACTCCCTCCAT
 TTGCCACGCCCCGGGAGGAATACTTTGTGGGAAGGAGAACGTGGAACATGTGACTCTGTACTCAGTGCACCTG
 CCGACGCGGACGGGTGCTGTGTGAGACAGAGGTGTGCCACCGCTGCTCTGCCAGAACCCTCAGCAGCCAGGA
 TTCCTGCTGCCACAGTGTACAGATCAACCTTTTCGGCCTTCTTGTCCCGCAATAACAGCGTACCTAATTACTG
 CAAAATGATGAAGGGGATATATTCTTGGCAGCTGAGTCTTGGAGCCTGACGTTTGTACCAGCTGCATCTGCAT
 TGATAGCGTAAATTAGCTGTTTCTCTGAGTCTTGCCTTCTGTATCCTGTGAAAGACCTGTCTTGAGAAAGGCCA
 GTGTTGTCTCTCATGCATAGAAGACAAATTCAAAGAAGTGGTGTGCCACTCAGTGGGAAGGCCTATGCCGA
 CGAGGAGCGTGGGACCTTGCAGCTGCACCCACTGTACTGCTGCGAGGCGAGACCTCTGCTCGACCGTCAG
 CTGCCCCCTCTGCGCTGTGTTGAGCCCATCAACGTGGAAGGAAGTGTGTCGCCAATGTGTCCAGAAATGTATGT
 CCCAGAACCACCAATATACCCATTGAGAAGACAAACCATCGAGGAGAGGTTGACCTGGAGGTTCCCCTGTGGCC
 CAGTGCATAGTGAATAATGATATCGTCCATCTCCCTTAGAGATATGGGTCACTCCAGGTAGATTACAGAGATAACAG
 CCGTCAACCAAGTGAAGATTCTTCACTGGACTTCCATGGCTCAGTTGTGGTCCCATAATTATATGCCTCTAT
 TATAATAGCATTTCTATTATCATCATCAGAAGAAACAGTGGATACCACCTGCTTGTGTTGTGATCGAACCAACTAA
 GCCTTCTTCTTAAATAATCAGCTAGTATCTGTGGACTGCAAGAAAGGAACAGAGTCCAGGTGGACAGTTCCTCA
 GAGAATGCTAAGAATTGCAGAACCAGATGCAAGATTAGTGGCTTCTACAGCATGCAAAAACAGAACCATCTACA
 GGCAGACAAATTTCTACCAAACAGTGTGAGAAAGGCAACTAGGATGAGGTTTCAAAGACGGAAGCAGCATAAAT
 CTGCTCTAAAAAGTAAACTAGAATTTGTGCACTTGTCTAGTGATTGTATTGGATTGTGACTTGTATGTACAGCGC
 TAAGACCTTACTGGGATGGGCTCTGTCTACAGCAATGTGCGAAGCAAGCATCCCACCTTTTCTCAAAAAA

142/550

FIGURE 142

MYLVAGDRGLAGCGHLLVSLGLLLLLLARSSTRALVCLPCDESKCEEPRNCPGSIVQGVCGCC
 YTCASQRNESCGGTFGIYGTCDRGLRCVIRPPLNGDSLTEYEAGVCEDENWTDDQLLGFKPCN
 ENLIAGCNIINGKCECNTIRTCNPFEPFSDMCLSAKRIEEKPDCKARCEVQFSPRCPE
 DSVLIEGYAPPGECCPLPSRCVCNPAGCLRKVCQPGNLNLVSKASGKPGECDDLYECKPVFG
 VDCRTVECPVPVQQTACPPDSYETQVRLTADGCCTLPTRCECLSGLCGFPVCEVGSTPRIVSRG
 DGTGPKCCDVFEVCVNDTKPACVFNNVEYYDGMFRMDNCRFCRCQGGVAICFTAQCGEINCER
 YYVPEGECCPVCEDPVYPFNNPAGCYANGLILAHGDRWREDDCTFCQCVNGERHCVATVCGQT
 CTNPVKVPGECCPVCEPTIITVDPPACGELSNCTLTGKDCINGFKRDHNGCRTCQCINTEEL
 CSERKQGCTLNCPPGFLTDAQNCEICECRPRPKKCRPIICDKYCPLGLLKNKHGCDICRCKK
 PELSCSKICPLGFGQDQSHGCLICKCREASAGAPPILSGTCLTVDGHGHHKNEESWHDGCRECY
 CLNGREMCALITCPVPACGNPTIHPGQCCPSCADDFVQKPELSTPSICHAPGGGEYFVEGETW
 NIDSCTQCTCHSGRVLCETEVCPPLLQNPSRTQDSCCPQCTDQPFPSLSRNNSVPNYCKND
 EGDIFLAAESWKPDVCTSCICIDSVISCFSESCPSVSCERPVLKRGQCCPYCIEDTIPKKVVC
 HFSGKAYADEERWDLDSCTHCYCLQGQTLCTVSCPPPLPCVEPINVEGSCCPMCPEMYVPEPT
 NPIEKTNRGEVDLEVPLWPTPSENDIVHLPRDMGHLQVDYRDNRHLHPSEDSSLDSIASVVV
 PIIICLSIIIAFLFINQKKQWIPLLCWYRTPTKPSLNNQLVSVDCCKGTRVQVDSSQRLRI
 AEPDARFSGFYSMQKQNLQADNFYQTV

Important features of the protein:**Signal peptide:**

amino acids 1-34

Transmembrane domain:

amino acids 940-962

N-glycosylation sites.

amino acids 71-75, 113-117, 330-334, 474-478, 746-750

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 992-996

N-myristoylation site.

amino acids 9-15, 58-64, 61-67, 75-81, 79-85, 362-368, 402-408, 407-413,
 439-445, 492-498, 511-517, 551-557, 558-564, 586-592, 606-612, 625-631,
 845-851

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 52-63, 844-855

Cell attachment sequence.

amino acids 314-317

Leucine zipper pattern.

amino acids 3-25

Eukaryotic thiol (cysteine) proteases cysteine active site.

amino acids 57-69

VWFC domain proteins.

amino acids 448-456, 382-390

C-terminal cystine knot proteins

amino acids 60-86

143/550

FIGURE 143

[illegible]

144/550

FIGURE 144

MVARVGLLLRALQLLLWGHLDAQPAERGGQELRKEAEAFLEKYGYLNEQVPKAPTSTRFSDAI
RAFQWVSQLPVSGVLDRA TL RQMTRPRCGVTD TNSYA AWAERISDLFARHRTKMRRKKRFAKQ
GNKWKQHL SYRLVNWPEHLPEPAVRGAVRAAFQLWSNVSALEFWEAPATGPADIRLTFFQGD
HNDGLGNAFDGPGGALAHAF LPRRGEAHFDQDERWSLSRRRGRNLFVFLAHEIGHTLGLTHSP
APRALMAPYYKRLGRDALLSWDDVLAVQSLYGKPLGGSVAVQLPGKLFTDFETWDSYSPQGRR
PETQGPKYCHSSFDAITVDRQQQLYIFKGSHFWEVAADGNVSEPRPLQERWVGLPPNIEAAAV
SLNDGDFYFFKGGRCWRFRGPKPVWGLPQLCRAGGLPRHPDAALFFPPLRRLILEKGARYYVL
ARGGLQVEPYYP RSLQDWGGIPEEVSGALPRPDGSIIFRDDR YWRLDQAKLQATTSGRWATE
LPWMGCWHANSGSALF

Important features of the protein:**Signal peptide:**

amino acids 1-22

N-glycosylation sites.

amino acids 164-168, 355-359

N-myristoylation sites.amino acids 92-98, 153-159, 193-199, 202-208, 288-294, 368-374,
509-515**Amidation site.**

amino acids 312-316

Neutral zinc metalloproteinases, zinc-binding region signature.

amino acids 237-247

Matrixins cysteine switch

amino acids 231-262, 271-284

Hemopexin domain protein

amino acids 66-108, 231-262

145/550

FIGURE 145

CGCGGCTAGGGCGCCGGAGCCGCACGCAGCCGCGGGGCTCCGAGAGGCGCGCACTGGGGCTGGGACTGCGCGGCG
CCGCCGCTGCGAGCGCCACTGAGCGGTGCGCGCAACTTCGGAGGCACAGCGCCGAGCCAGGCGAGCGCTCAGAGA
CCCGGAGCCAGAGGGGCGCGCCGGAGCGCTCGTTCCGAGAGCCGGCGCCAGGCACCCACCGCGCTCCGAGTGCCAGG
CGGCCCTCCGCGCAGCGTGGCTTCCGCTGCCCCACGGAAAGGCACGGGCTGGCGCTGCCGGGCGCCGGGGAGGAC
GGCGAGGAGGAGGCGGCGCGCGGAGACGGCGCGCGGAGACTGGGGCCAGGGAGACAGCCCTGGGGGAGAGGC
GCCCCAACCAGGCGCGGGAGC**ATG**GGGGCCCCGGAGCGGAGCTCGGGGCGCGCTGCTGCTGGCACTGCTGCTCTG
CTGGGACCCGAGGCTGAGCCAAGCAGGCACTGATTCTGGCAGCGAGGTGCTCCCTGACTCCTTCCCGTCAGCGCC
AGCAGAGCCGCTGCCCTACTTCCCTGCAGGAGCCACAGGACGCCCTACATTGTGAAGAACAAGCCTGTGGAGTCCG
CTGCCGCGCCTTCCCCGCCACACAGATCTACTTCAAGTGCAACGGCGAGTGGGTGAGCCAGAACGACCACGTCAC
ACAGGAAGCCCTGGATGAGGCCACCGGCTGCGGGTGC GCGAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGA
GGAGCTCTTTGGGCTGGAGGATTACTGGTGCCAGTGCGTGGCCTGGAGCTCCGCGAGGCACCAAGAGTCGCGG
AGCCTACGTCCGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTGGGCAAGGAGGTGCCCTGGACCA
TGAGGTTCTCTGCAGTGCCGCCCGCCGGAGGGGGTGCCTGTGGCCGAGGTGGAATGGCTCAAGAATGAGGATGT
CATCGACCCCAACCCAGGACACCAACTTCCCTGCTCACCATCGACCACAACCTCATCATCCGCCAGGCGCCGCTGTC
GGACACTGCCAACTATACCTGCGTGGCCAAAGAACTCGTGGCCAAACGCCGAGCACCCTGCCACCGTCATCGT
CTACGTGAATGGCGGCTGCTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAACCGCTGTGGCCGAGGCTGGCAGAA
GCGCACCCGGACCTGCACCAACCCCGCTCCACTCAACGGAGGGGCTTCTGCGAGGGCCAGGCATTCCAGAAGAC
CGCCTGCACCAACCATCTGCCAGTCGATGGGGCGTGGACGGAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTG
TGCCCACTGGCGTAGCCGCGAGTGATGGCGCCCCACCCAGAACGGAGGCGGTGACTGCAGCGGGACGCTGCT
CGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAAATAAGAAACTCTAAGCGACCCCAACAGCCACCTGCT
GGAGGCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCTGTGGTCTGTTGCAATCCTCATGGC
GGTGGGGTGGTGGTGTACCGCCGCAACTGCCGTGACTTCGACACAGACATCACTGACTCATCTGCTGCCCTGAC
TGGTGGTTTCCACCCCGTCAACTTTAAGACGGCAAGGCCAGCAACCCGAGCTCCTACACCCCTCTGTGCCTCC
TGACCTGACAGCCAGCGCCGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAAATCCCCAT
GACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGCACCACGGGCTCTGG
GCCAGGCTGGCAGATGGGGCTGACCTGCTGGGGTCTTGCCGCTGGCACATACCTTAGCGATTTGCCCGGGGA
CACCCTTCTGCACTGCGCAGCGCCAGCCTCGGTTCCAGCAGCTCTTGGGCTGCCCGAGACCCAGGGAG
CAGCGTCAGCGGCACCTTTGGCTGCTGGTGGGAGGCTCAGCATCCCCGGCACAGGGGTGAGCTGTGCTGGTCC
CAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAGGCAGAAAGTACCCTCCCGCT
TTCAGAAAGGACCCAGACAGTATTGAGCCCCCTCGGTGACCTGTGGACCCACAGGCCTCCTGCTGTGCCGCCCCGT
CATCCTCACCATGCCCCACTGTGCCGAAGTCAGTGCCCGTGAAGTGGATCTTTAGCTCAAGACCCAGGCCACCA
GGGCACTGGGAGGAGTGGTGACCTGGATGAGGAGACCTGAACACACCCTGCTACTGCCAGTGGAGCCAG
GGCCTGTACATCCTGTGGACCACTGGGCACCTAGCTGTTACCGGGCAGTCCCTATTTCCCGCTCAGCAGTCAA
GCGGCTCAGCTGGCGCTTTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGA
GGACACGCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAA
ACCGCTAATGTTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCATGCCATTTGGAGGAG
CAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATACATTTGGAGTGGCAGCCAGAAGGCCCTCCACTGCAC
TTTACCCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTCACCTGCAAGATCTGCGTGC GGCAAGTGGAAAGGGGA
GGGCCAGATATTCCAGCTGCATACCACTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCC
TGGCAGCACTGTACCAACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATATGCAA
CAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAGCTCTCTATGGACCGGTACCT
GAATTACTTTGCCACCAAGCGAGCCCCACGGGTGTGATCCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGG
GGACCTCAACAGCCTGGCGAGTGCCCTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGG
GGACTGCT**GAG**CCCTCCTGGGACAGCGGGCTGGCAGGAGTGGCAGGAGGCAGGTGCAGGGAGGCTGGGGCAGCC
TCCTGATGGGGATGTTGGCCTCTGCTTCCCTCCAGTTACAGCCAGAGTTGCCCTCCTCCTCCTCTTCCCCAA
CCCCAGACCATGACCAGCCTTAGAAAATCCATGTACTCTGTTGTTAGAGGGCCAGAGTTCCTTCTCCACCCCC
GCTCTCTCTCTTGGCCTGAGATCTCTGTGCAGGAACCAAGATGGGGCTGAAGCCTCTGGAGGAGTTGGTTGG
GGCGGGCAGGCAGGAGGCCCTCCCTCCACCCCCACCCCTCAGCCCCGCAACTTCTGGGTTCCGTGGGTTTGTAG
TTCCGTTCTTCTGTTTCTTCTCCTCGTTATTGATTTCTCCTTTCTCCCTAAGCCCCCTTCTGCTTCCACGCCCTTT
TCCTCTTTGAAGAGTCAAGTACAATTAGACAAACTGCTTTCTCCTGTCCAAAAGCAAAAAGGCAAGGAAAGAA
AGAAAGCTTCAGACCGCTAGTAAGGCTCAAGAAGAAGAAAAACACCAAAACCACAAGGGAAAAGAAAAACCCAG
TTTCTTAGGAAACGCAACGATTATTATCCAGATTATTGGATAAGTCCTTTTAAAA

146/550

FIGURE 146

MGARSGARGALLLALLLCWDPRLSQAGTDSGSEVLPDSEFPSAPAEPLPYFLQEPQDAYIVKNK
 PVELRCRAFPATQIYFKCNGEWSQNDHVTQEGLEATGLRVREVQIEVSRQQVEELFGLEDY
 WCQCVAWSSAGTTKSRRAYVRIAYLRKNFDQEPLGKEVPLDHEVLLQCRPPEGVPVAEVEWLK
 NEDVIDPTQDTNFLTIDHNLIIRQARLSDTANYTCVAKNIVAKRRSTTATVIVYVNGGWSSW
 AEWSPCSNRCGRGWQKRTRCTNPAPLNGGAFCEGQAFQKTACTTICPVDGAWTEWSKWSACS
 TECAHWRSRECMAPPPQNGGRDCSGTLLDSKNCTDGLCMQNKKTLSDPNSHLLEASGDAALYA
 GLVVAIFVVVAILMAVGVVYRRNCRDFTDITDSSAALTGGFHPVNFKTARPSNPQLLHPSV
 PPDLTASAGIYRGPVYALQDSTDKIPMTNSPLLDPLPSLKVKVYSSSTTGSGPGLADGADLLG
 VLPPGTYPSTDFARDTHFLHLRSASLGSQLGLPRDPGSSVSGTFGCLGGRLSIPGTGVSLLV
 PNGAIPQKGFYEMYLLINKAESTLPLSEGTQTVLSPSVTCGPTGLLLCRPVILTMPHCAEVSA
 RDWIFQLKTAHQGHWEVVTLDEETLNTPCYQCLEPRACHILLDQLGTYVFTGESYSRSAVK
 RLQLAVFAPALCTSLEYSLRVYCLEDTPVALKEVLELERTLGGYLVVEPKPLMFKDSYHNRL
 SLHDLPHAHWRSKLLAKYQEIIFYHIWGSQKALHCTFTLERHSLASTEELTCKICVRQVEGEG
 QIFQLHTTLAETPAGSLDTLCSAPGSTVTTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRM
 LAQKLSMDRYLNYFATKASPTGVILDLWEALQQDDGDLNSLASALEEMGKSEMLVAVATDGDC

Important features of the protein:**Signal peptide:**

amino acids 1-26

Transmembrane domain:

amino acids 374-395

N-glycosylation sites.

amino acids 222-225, 347-350

Glycosaminoglycan attachment site.

amino acids 492-495

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 233-236, 234-237

Casein kinase II phosphorylation sites.

amino acids 30-33, 87-90, 251-254, 341-344, 359-362, 629-632, 651-654, 706-709, 757-760, 827-830, 925-928, 941-944

Tyrosine kinase phosphorylation sites.

amino acids 216-223, 773-780

N-myristoylation sites.

amino acids 2-7, 6-11, 27-32, 96-101, 137-142, 179-184, 247-252, 281-286, 334-339, 379-384, 491-496, 495-500, 509-514, 542-547, 547-552, 550-555, 553-558, 560-565, 611-616, 785-790, 834-839, 844-849

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 541-551

ATP/GTP-binding site motif A (P-loop).

amino acids 926-933

Growth factor and cytokines receptors family signature 2.

amino acids 306-312

147/550

FIGURE 147

GAGAGGGACAGAGGCTGGAGAAGGATGTATGGCCTGCCCTGGGCTTGTCTGTTCCCTCCTGAGCCTGAGCCCCCTT
ACCTTCCTGACCCCATGAAGCACACACTGGCTCTGCTGGCTCCCCTGCTGGGCCTGGGCCTGGGCCTGGCCCTGA
GTCAGCTGGCTGCAGGGGCCACAGACTGCAAGTTCTTGGCCCCGGCAGAGCACCTGACATTACCCCCAGCAGCCA
GGGCCCCGGTGGCTGGCCCCCTCGAGTTCGTGCGCCAGGACTCCTGGACTCCCTCTATGGCACCGTGCGCCGCTTCC
TCTCGGTGGTGCAGCTCAATCCTTTCCCTTCAGAGTTGGTAAAGGCCCTACTGAATGAGCTGGCCTCCGTGAAGG
TGAATGAGGTGGTGGGTACGAGGCGGGCTACGTGGTATGCGCTGTGATCGCGGGGCTTACCTGCTGCTGGTGC
CCACTGCCGGGCTTTGCTTCTGCTGCTGCCGCTGCCACCGGCGCTGCGGGGACGAGTGAAGACAGAGCACAAAG
CGCTGGCCTGTGAGCGCGCGGCCCTCATGGTCTTCCCTGCTGCTGACCACCTCTTGCTGCTGATTGGTGTGGTCT
GTGCCTTTGTACCAACCAGCGCACGCATGAACAGATGGGCCCCAGCATCGAGGCCATGCCTGAGACCTGCTCA
GCCTCTGGGGCTGGTCTCTGATGTCCCCAAGAGCTGCAGGCCGTGGCACAGCAATTCTCCCTGCCCCAGGAGC
AAGTCTCAGAGGAGCTGGATGGTGTGGTGTGAGCATTGGGAGCGCGATCCACACTCAGCTCAGGAGCTCCGTGT
ACCCCTTGCTGGCGGCGGTGGGCAGTTTGGGCCAGGTCTGCAGGTCTCCGTGCACCACCTGCAAACCTTGAATG
CTACAGTGGTAGAGCTGCAGGCCGGGCAGCAGGACTGGAGCCAGCCATCCGGGAACACCGGGACCGCTCCTTG
AGCTGCTGCAGGAGGCCAGGTGCCAGGGAGATTGTGCAGGGGCCCTGAGCTGGGCCCCGACCCTGGAGCTGGGTG
CTGACTTCAGCCAGGTGCCCTCTGTGGACCATGTCTGCACCAGCTAAAAGGTGTCCCGAGGCCAACTTCTCCA
GCATGGTCCAGGAGGAGAACAGCACCTTCAACGCCCTTCCAGCCCTGGCTGCCATGCAGACATCCAGCGTGGTGC
AAGAGCTGAAGAAGGCAGTGGCCCAGCAGCCGAAGGGGTGAGGACACTGGCTGAAGGGTTCCCGGGCTTGGAGG
CAGCTTCCCGCTGGGCCAGGCACTGCAGGAGGTGGAGGAGAGCAGCCGCCCCCTACCTGCAGGAGGTGCAGAGAT
ACGAGACCTACAGGTGGATCGTGGGCTGCGTGTGCTCCGTGGTCTTATTCGTGGTGTCTGCAACCTGCTGG
GCCTCAATCTGGGCATCTGGGGCTGTCTGCCAGGGACGACCCAGCCACCCAGAAGCCAAGGGCGAGGCTGGAG
CCCGCTTCCCTCATGGCAGGTGTGGGCCTCAGCTTCCCTCTTGTGTCACCCCTCATCTCCTGGTGTTCGCCACCT
TCCTGGTGGGTGGCAACGTGCAGACGCTGGTGTGCCGGAGCTGGGAGAACGGCGAGCTCTTTGAGTTTGACAGCA
CCCCAGGGAACCTGCCCCCTCCATGAACCTGTCGCAACTTCTTGGCCTGAGGAAGAACATCAGCATCCACCAAG
CCTATCAGCAGTGCAAGGAAGGGGCAGCGCTCTGGACAGTCTGCAGCTCAACGACTCCTACGACCTGGAGGAGC
ACCTGGATATCAACCAGTATACCAACAAGCTACGGCAGGAGTTGCAGAGCCTGAAAGTAGACACACAGAGCCTGG
ACCTGCTGAGCTCAGCCGCCCGCCGGGACCTGGAGGCCCTGCAGAGCAGTGGGCTTACGCGCATCCACTACCCCG
ACTTCTCGTTCAGATCCAGAGGCCGCTGGTGAAGACCAGCATGGAGCAGCTGGCCCAGGAGCTGCAAGGACTGG
CCCAGGCCCAAGACAATTCTGTGCTGGGGCAGCGGCTGCAGGAGGAGGCCCAAGGACTCAGAAACCTTCAACAGG
AGAAGGTCGTCCCCCAGCAGAGCCTTGTGGCAAAGCTCAACCTCAGCGTCAGGGCCCTGGAGTCTCTGCCCCGA
ATCTCCAGCTGGAGACCTCAGATGTCTAGCCAATGTACCTACCTGAAAGGAGAGCTGCCTGCCTGGGCAGCCA
GGATCCTGAGGAATGTGAGTGTGTTTCTGGCCCCGGGAGATGGGCTACTTCTCCAGTACGTGGCCTGGGTGA
GAGAGGAGGTGACTCAGCGCATTGCCACCTGCCAGCCCCCTCTCCGAGCCCTGGACAACAGCCGTGTGATCCTGT
GTGACATGATGGCTGACCCCTGGAATGCCTTCTGGTTCTGCCTGGCATGGTGCACCTTCTTCTGATCCCCAGCA
TCATCTTTGCCGTCAAGACCTCCAAATACTTCCGTCTATCCGGAACGCCTCAGCTCCACCAGCTCTGAGGAGA
CTCAGCTCTTCCACATCCCCCGGTTACCTCCTGAAGCTGTAGGGCCTTGTGGGGTGAGGTGACCTGAGGCTG
CCTGTCTCTCCCTTTGATTTAGCCTGGGCCACAGGACTTCGGTAGCTCTTGCCCCAGAGCCAGGCTGGCATCCA
GGCCTGGACTGTCCCACTTCCGGCTTACCTGGCCCCACCTTGCTGCTCCTTCCACCCCTTTTGTCTACGAC
CCCCATCATTACGCTCAGAATCACATGGGACTTCTGTGCAGCTGCAGAGCCAGCAAGTCCCTACAGGTGTCAAC
CGTTACCCCCATGCTGGTGGCATCCTCACAGGAAGAGCCTGTTCTCCACCTGCTGGAGCCTGGACCTGGGGTGG
GACAGAGGCCTCGTCAACCCCACTCCCTTCCCGTGTGTCTCCCCCTGCCAAGCCTCCCCCTGCCAAGCCTCC
CCCTGCCCTCTCTGAGCCCTCGCCCCCACACCGTCTCATCTGGCCTCCCCCTGGCCCCACTTCCCTCTT
ATGCCCTTCTGGCCCTTTGCTTCTCCTTAGTCCCCCTTACCATATCTCCACTGCTACCTTGCTGGCCCCA
GAGACCACCTGCCCAACCAACCACTCAGGTAACGCCACTAATCAGGCAGGGGCCACCATGGCCTAGGTCTGGG
CTGGCTGCAGGCCCTGCCTCATGGCCTCTGAGCCCTCCACTGCCCCAGGGCCTTGGGCCCTCTGCAGATCTCATC
CAGGATTTATTGTTGCCAGTGGGGTGAGGGAGGCCTGTCTGAAGGCCGAGCCTCCCTGCCTGCACCCAAGTTAG
AAATGGGGGTACCAGCACTTAGCTTCTCTGAGTGTCTGGCTCCCAAGGAAGGGACCTGGGACCTGGGCCACAGT
GGGGGCTTGCCCTTACCTCTTCAAGGAAGCATCTTCCACAGCCCCCACCCTTCTTAGGAGTGTATCTGGT
GGCCAGAAGGATTTTGCACGGCCCTTTTATCCTGCGCATGTGGCCTAGGGTCATCCCCAGGCCATCCCTGTG
TCAGCCCTGAGTGTGGACACTGCGTTCAGAAATGAGGAAGAGGAGAGAGAAGAGATGGACAGACCTCAGATCC
ATTAAAGTGTCTCACTTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

148/550

FIGURE 148

MKHTLALLAPLLGLGLGLALSQLAAGATDCKFLGPAEHLTFTPAARARWLAPRVRAPGLL
DSLYGTVRRFLSVVQLNPFPSSELVKALLNELASVKVNEVVRYEAGYVVCVAVIAGLYLLL
PTAGLCFCCCRCHRRCCGRVKTEHKALACERAALMVFLLLTLLLLIGVVCAFTVNQORTH
EQMGPSIEAMPETLLSLWGLVSDVPQELQAVAQQFSLPQEQVSEELDGVGVSIGSAIHTQ
LRSSVYPLLAAGVSLGQVLQVSVHHLQTLNATVVELQAGQQDLEPAIREHRDRLELLQE
ARCQGDCAALSWARTLELGADFSQVPSVDHVLHQLKGVPEANFSSMVQEENSTFNALPA
LAAMQTSSVQELKKAVAQQPEGVRTLAEGFPGLEAASRWAQALQEVESRPPYLQEVQR
YETRWIVGCVLCSVVLFFVLCNLLGLNLGIWGLSARDDPSHPEAKGEAGARTLMAGVGL
SFLFAAPLILLVFATFLVGGNVQTLVCRSWENGELFEFADTPGNLPPSMNLSQLLGLRKN
ISIHQAYQQCKEGAALWTVLQLNDSYDLEHLDINQYTNKLRQELQSLKVDTSQSLDLLSS
AARRDLEALQSSGLQRIHPDFLVQIQRPVVKTSMEQLAQELQGLAQADNSVLGQRLQE
EAQGLRNLHQEKVVPQQSLVAKLNLVRALESSAPNLQLETSDVLANVTYKLGELPAWAA
RILRNVSECFLEMGYFSQYVAWVREEVTQRIATCQPLSGALDNSRVILCDMMADPWNA
FWFCLAWCTFFLIPSIIFAVKTSKYFRPIRKRLSSTSSEETQLFHIPRVTSCLK

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 105-125, 153-173, 428-449, 476-500, 778-797

N-glycosylation sites:amino acids 270-273, 343-347, 352-356, 530-534, 540-546, 563-567,
684-688, 707-711, 725-729**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 811-815

Tyrosine kinase phosphorylation site.

amino acids 95-103

N-myristoylation sites.amino acids 13-19, 15-21, 17-23, 26-32, 58-64, 124-130, 168-174,
228-234, 230-236, 320-326, 338-344, 393-399, 429-435, 446-452,
477-483, 500-506, 536-542, 644-650, 761-767**Phospholipase A2 histidine active site.**

amino acids 129-137

4Fe-4S ferredoxins, iron-sulfur binding region signature.

amino acids 126-138

Mitochondrial energy transfer proteins signature.

amino acids 80-89

149/550

FIGURE 149

CACAGCTCCCTTCCCAGGACGTGAAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCTTTCCA
GCCTGCTCTGTATCCTGCTTCTCTGCTTCTCCATCTTCTCCACAGAAGGGAAGAGGCGTCCTG
CCAAGGCCTGGTCAGGCAGGAGAACCAGGCTCTGCTGCCACCGAGTCCCTAGCCCCAACTCAA
CAAACCTGAAAGGACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCCGCC
TTTGGGTGGTGCCTGGGGCACTCCCACAGGTGTAGCACTCCCAAAGCAAGACTCCAGACAGCG
GAGAACCTCATGCCTGGCACCTGAGGTACCCAGCAGCCTCCTGTCTCCCCTTTCAGCCTTCAC
AGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCTGGGAAAGTTCCAG
AACTCCACGTCCTTGTCTCAATTGTGCCATCAACTTTCAGAGCTATCATGAGCCAACCTCACC
CCACAGGGCCTCAGTCGCCACCATGTGGGCCTCTCCAGTGCAAACCACCGAGCATTCACCAT
GACCGGTCACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAAGGTGGCAAGCACC
CAAGGGTGGCTGACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTCCATTACGCCTCCTGGCA
TTTAACTACCAGCATCCAGTGGTCCCCAAGGAATCCCTTCCTAGCCTCCTGACATGAGTCTGC
TGGAAGAGCATCCAAACAAACAAGTAATAAATAAATAAATAAACTCA

150/550

FIGURE 150

MRLVLSSLLCILLLCFSIFSTEGKRRPAKAWSGRRTRLCCHRVSPNSTNLKGHHVRLCKPC
KLEPEPRLWVVP GALPQV

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 48-52

Amidation sites.

amino acids 23-27, 33-37

151/550

FIGURE 151

CACCGGAGGGGCACGCAGCTGACGGAGCTGCGCTGCGTTTCGCTCGTTTGCCTCGCGCCCTCCA
CTGGAGCTGTTTCGCGCCTCCCGGCTCCCACCGCAGCCCACCCGGCAGAGGAGTCGCTACCAGC
GCCCAGTGCCTCTGTTCAGTCCGCAAACCTCCTTGCCGCCCGCCCCGGGCTGGGCACCAAATAC
CAGGCTACCAATGGTCTACAAGACTCTCTTCGCTCTTTGCATCTTAACATGCAGGATGGAGGGTA
CAGAGTCTGCCTACATCAGCTCCTTTGTCTGTTTCTCTTCCGACAAACATTGTACCACCGACC
ACCATCTGGACTAGCTCTCCACAAAACACTGATGCAGACACTGCCTCCCCATCCAACGGCACT
CACAACAACCTCGGTGCTCCCAGTTACAGCATCAGCCCCAACATCTCTGCTTCCCTAAGAACATT
TCCATAGAGTCCAGAGAAGAGGAGATCACCAGCCCAGGTTTGAATTGGGAAGGCACAAACACA
GACCCCTCACCTTCTGGGTTCTCGTCAACAAGCGGTGGAGTCCACTTAACAACCACGTTGGAG
GAACACAGCTCGGGCACTCCTGAAGCAGGCGTGGCAGCTACACTGTCGCAGTCCGCTGCTGAG
CCTCCACACTCATCTCCCCTCAAGCTCCAGCCTCATCACCTCATCCCTATCAACCTCACCA
CCTGAGGTCTTTTCTGCCTCCGTTACTACCAACCATAGCTCCACTGTGACCAGCACCCCAACCC
ACTGGAGCTCCAACATGCACCAGAGTCCCCGACAGAGGAGTCCAGCTCTGACCACACACCCACT
TCACATGCCACAGCTGAGCCAGTGCCCCAGGAGAAAACACCCCCAACAACTGTGTGAGGCAAA
GTAGAACATGCATTAAAGTTCAGGCAGCATCGCCGCCATTACCGTGACAGTCATTGCCGTGGTG
CTGCTGGTGTGTTGGAGTTGCAGCCTACCTAAAAATCAGGCATTCTCCTATGGAAGACTTTTG
GACGACCATGACTACGGGTCCTGGGGAAACTACAACAACCCTCTGTACGATGACTCCTAACAA
TGGAATATGGCCTGGGATGAGGATTAAGTGTCTTTATTTATAAGTGCTTATCCAGTAGAATT
AATAAGTACCTGATGCGCATTGAACGACAATCTTAAGCCCTGTTTTGTTGGTATGGTTGTTTT
TGTTTTCTCTCCCTCTCCTCTGGCTGCTACAACCTTCCCCTTTCTGGTACAAGAAGAACCATTCT
TTAAAGGTGAGTGGAGGCTGATTTGCAGCTGAAGTGGGCCAGCCTTGACCAGCCAGGCCAGA
CCACCATGGTGAAGGCTTCTTTCCCCACTGCAGGACCCACTTTGAGAAGGATCGAGGAGGAGG
ATTTGGGTTGTTTTGTTAGGGGTTACTTTAGGGGAAACATTTTCAATTTGTGTTATTTCTTAAAC
TTCTATTTAGGAAATTACATTAAGTATTAATGAGGGGAAAGGAAATGAGCTCTACGAGGATTT
CACCTTGCTGGGAGAGAGCAGGGTTTTCTCAGATTCCTTTTTTAATCTCTATTTATCTGGTTG
TTTCTGACAGGATGCTGCCTGCTTGGCTCTACGAGCTGGAAAGCAGCTTCTTAGCTGCCTAAT
TAATGAAAGATGAAAATAGGAAGTGCCCTGGAGGGGGCCAGCAGGTCACGGGGCAGAATCTCT
CAGGTTGCTGTGGGATCTCAGTGTGCCCTACCTGTTCTCCCCTCCAGGCCACCTGTCTCTGT
AAAGGATGTCTGCTCTGTTCAAAGGCAGCTGGGATCCCAGCCCACAAGTGATCAGCAGAGTT
GCATTTCCAAAGAAAAAGGCTATGAGATGAGCTGAGTTATAGAGAGAAAGGGAGAGGCATGTA
CGGTGTGGGGAAGTGGAAGAGAAGCTGGCGGGGGAGAAGGAGGCTAACCTGCACTGAGTACTT
CATTAGGACAAGTGAGAATCAGCTATTGATAATGGCCAGAGATATCCACAGCTTGGAGGAGCC
CAGAGACTGTTTGCTTTATACCCACACAGCAACTGGTCCACTGCTTTACTGTCTGTTGGATAA
TGCTGTAAATGTTTAAAAAC

152/550

FIGURE 152

MVYKTLFALCILTAGWRVQSLPTSAPLSVSLPTNIVPPTTIWTSSPQNTDADTASPSNGTHNN
SVLPVTASAPTSLLPKNISIESREEEITSPGSNWEGTNTDPSPSGFSSTSGGVHLTTTLEEHS
SGTPEAGVAATLSQSAAEPPTLISPQAPASSPSSLSTSPPEVFSASVTTNHSSTVTSTQPTGA
PTAPESPTTESSSDHTPTSHATAEPVPQEKTPPTTVSGKVMCELIDMETTTTTFPRVIMQEVEH
ALSSGSIAAITVTVIAVLLVFGVAAYLKIRHSSYGRLLDDHDYGSWGNYNPLYDDS

Important features of the protein:**Signal peptide:**

amino acids 1-20

Transmembrane domain:

amino acids 258-278

N-glycosylation sites.

amino acids 58-61, 62-65, 80-83, 176-179

Casein kinase II phosphorylation sites.

amino acids 49-52, 85-88, 95-98, 100-103, 120-123, 121-124, 141-144, 164-167, 191-194, 195-198, 200-203

Tyrosine kinase phosphorylation site.

amino acids 289-296

N-myristoylation sites.

amino acids 59-64, 115-120, 128-133, 133-138, 257-262, 297-302

153/550

FIGURE 153

[illegible]

154/550

FIGURE 154

MLVHCVGLLLTGALLGLTLGAGALLASEPIYQPPSAWVPAGGLVGLALLGALLTLRWPRPFTV
LGTTLGSAVLVACVDYFLEGLALGSLGQRLQTLPALPSLC

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 38-55, 60-78

N-myristoylation sites.

amino acids 7-13, 12-18, 16-22, 22-28, 41-47, 50-56, 84-90, 88-94

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 67-78

155/550

FIGURE 155

TGCAATTAAAGGAGTCGGGTCTCTAACTGTTGATCTGTTTTTTTCCCTTCTGAGCAATGGAGC
TTACCATCTTTATCCTGAGACTGGCCATTTACATCCTGACATTTCCCTTGTACCTGCTGAACT
TTCTGGGCTTGTGGAGCTGGATATGCAAAAAATGGTTCCCCTACTTCTTGGTGAGGTTCACTG
TGATATACAACGAACAGATGGCAAGCAAGAAGCGGGAGCTCTTCAGTAACCTGCAGGAGTTTG
CGGGCCCCCTCCGGGAAACTCTCCCTGCTGGAAGTGGGCTGTGGCACGGGGGCCAACTTCAAGT
TCTACCCACCTGGGTGCAGGGTGACCTGTATTGACCCCAACCCCAACTTTGAGAAGTTTTTGA
TCAAGAGCATTGCAGAGAACCGACACCTGCAGTTTGAGCGCTTTGTGGTAGCTGCCGGGGAGA
ACATGCACCAGGTGGCTGATGGCTCTGTGGATGTGGTGGTCTGCACCCTGGTGCTGTGCTCTG
TGAAGAACCAGGAGCGGATTCTCCGCGAGGTGTGCAGAGTGCTGAGACCGGGAGGGGCTTTCT
ATTTTCATGGAGCATGTGGCAGCTGAGTGTTGCACTTGAATTACTTCTGGCAACAAGTCCTGG
ATCCTGCCTGGCACCTTCTGTTTGATGGGTGCAACCTGACCAGAGAGAGCTGGAAGGCCCTGG
AGCGGGCCAGCTTCTCTAAGCTGAAGCTGCAGCACATCCAGGCCCCACTGTCTGGGAGTTGG
TGCGCCCTCATATCTATGGATATGCTGTGAAATTAGTGTGAGCTGGCAGTTAAGAGCTGAATGG
CTCAAAGAATTTAAAGCTTCAGTTTTACATTTAAAATGCTAAGTGGGAGAAGAGAAACCTTTT
TTTTGGGGGGCGGTTTTTTTTGGTTTGTGTTGGTTTTTTTTTTTTTTTTTTGGCAGGAGAATCTC
TTGAACCCAGAAGGCGAAGGTTGCAGTGAACCGAGATCATGCCATTGTACTCTAGCCTGGGTG
ACAAGAGCAAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAGAAGTAGAGACAGGGAGAC
GGGGTCTCACTGTGTTGCCTAGGCCGGTCTTGAACCTCCTGGGCTCAAGTGATTCTCCACCTT
GACCTCCTAAATTGTTGGGATTACAGGTGTGAGACAGTGCACCTGGCCGAAATAGCTCAAGTT
TCTGAAAAACAAATCTGAATCTATTTGTTATTCTTAGCGTCACTGGTCTGGCTTTCAGAATTA
ACATACAAGGTTGCCACACCTAGTTCTGCCAGCTTTATGTCTTTTATTCCAGTATTCCACCA
AAGTTTGTTCCTGCATTCCAGTTCTCAAGTCTTAAGATAAAGATTGTACTTGACAGTTTAG
TATATCCATAAACTATTTGAGGTGGTTAAGGTTCTTGGGTTCAATTTTCCTTAATACTTTGCT
GAATATTGTAGATTGTAGGCAATGAAAAAGTCTACTAAATTAGGAAAACCTTGAATAATTAGG
TATCCTAGGTAAGAGCCCCTAACATCAAGCAATCTGTGAGTCTGTAAAGAAATAAATATTTT
TTGGATTATTCTTATCTAATTCCACCCCTGTTGGAAGATGATTTCTTTGTTCTTTGCAACTAT
GGAAGCTGTGAAAATCATCACAAGTGCCTCTGAAAGCGAGTGTTAGGTTGGTTAGAGGGTTTA
ATATTTTCTGCAATGGTTTGTAGGAATTTTAATAAATGTAGTATATTTTCTGAGATGATTTTG
TAAAAGTACTATTTTAAATATCAAATCAACCAATAAATTCACATTTGTGTTAGGAACAAAA

156/550

FIGURE 156

MELTIFILRLAIYILTFPLYLLNFLGLWSWICKKWFYFLVRFTVIYNEQMASKKRELFNLQ
EFAGPSGKLSLLEVGCCTGANFKFYPPGCRVTCIDPNPNFEKFLIKSIAENRHLQFERFVAA
GENMHQVADGSVDVVCTLVLCVKNQERILREVCRLRPGGAFYFMEHVAAECSTWNYFWQQ
VLDPAWHLLFDGCNLTRESWKALERASFSLKLQHIQAPLSWELVRPHIYGYAVK

Signal peptide:

amino acids 1-29

N-glycosylation site.

amino acids 203-207

N-myristoylation sites.

amino acids 78-84, 80-86, 91-97, 201-207

157/550

FIGURE 157

CCGCTGAGATGTACGAACTTCCGGTTCTCCGGGCAGCTGCCACTGCTGTAGCTTCTGCCACCT
GCCACGACCGGGCCTCTCCCTGGCGTTTGGTCACCTCTGCTTCATTCTCCACCGCGCCTATGG
TCCCTCTTGGAGCCAGCGTGGCGGGCCTGGCGGGCTCCCGGGTGGTGAGAGAGCGGTCCGGGAA
CGATGAAGGCCTCGCAGTGCTGCTGCTGTCTCAGCCACCTCTTGGCTTCCGTCCTCCTCCTGC
TGTTGCTGCCTGAACTAAGCGGGCCCCCTGGCAGTCCTGCTGCAGGCAGCCGAGGCCGCGCCAG
GTCTTGGGCCTCCTGACCCTAGACCACGGACATTACCGCCGCTGCCACCGGGCCCTACCCCTG
CCCAGCAGCCGGGCCGTGGTCTGGCTGAAGCTGCGGGGCCGCGGGGCTCCGAGGGAGGCAATG
GCAGCAACCCTGTGGCCGGGCTTGAGACGGACGATCACGGAGGGAAGGCCGGGGAAGGCTCGG
TGGGTGGCGGCCTTGCTGTGAGCCCCAACCTGGCGACAAGCCCATGACCCAGCGGGCCCTGA
CCGTGTTGATGGTGGTGAGCGGCGCGGTGCTGGTGTACTTCGTGGTCAGGACGGTCAGGATGA
GAAGAAGAAACCGAAAGACTAGGAGATATGGAGTTTGGACACTAACATAGAAAATATGGAAT
TGACACCTTTAGAACAGGATGATGAGGATGATGACAACACGTTGTTTGATGCCAATCATCCTC
GAAGATTAGAATGTGCCTTTTGATGAAAGAACTTTATCTTTCTACAATGAAGAGTGGAATTC
TATGTTTAAGGAATAAGAAGCCACTATATCAATGTTGGGGGGGTATTTAAGTTACATATATTT
TAACAACCTTTAATTTGCTGTTGCAATAAATACCGTATCCTTTTATTATATCTTTATATGTAT
AGAAGTACTCTATTAATGGGCTCAGAGATGTTGGGGATAAAGTATACTGTAATAATTTATCTG
TTTGAAAATTACTATAAAACGGTGTTTTCTGGTCGGTTTTTGTTCCTGCTTACCATATGATT
GTAAATTGTTTTATGTATTAATCAGTTAATGCTAATTATTTTTGCTGATGTCATATGTTAAAG
AGCTATAAATTCCAACAACCAACTGGTGTGTAAAAATAATTTAAATTTCTTTACTGAAAGG
TATTTCCCATTTTTGTGGGGAAAAGAAGCCAAATTTATTACTTTGTGTTGGGGTTTTTAAAT
ATTAAGAAATGTCTAAGTTATTGTTTGCAAAACAATAAATATGATTTTAAATTCTCTTAAAA
AAAA

158/550

FIGURE 158

MKASQCCCCLSHLLASVLLLLLLPELSGPLAVLLQAAEAAPGLGPPDPRPRTLPPPLPPGPTPA
QQPGRGLAEAAGPRGSEGGNGSNPVAGLETDDHGGKAGEGSVGGGLAVSPNPGDKPMTQRALT
VLMVVSGAVLVYFVVRTVRMRRNRKTRRYGVLDTNIENMELTPLEQDDEDDNTLFDANHPRR

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 124-140

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.amino acids 69-75, 78-84, 81-87, 97-103, 103-109, 106-112,
157-160

159/550

FIGURE 159

GCTGCAGGCGGGGACGGCTACACCA**ATG**GGCGGGCTGCTGCGGGCCGCCGGCTGCCGCCGCTG
CTTTCGCCGCTGCTGCTTCTGCTGGTTGGGGGAGCGTTCCTGGGTGCCTGTGTGGCTGGGTCT
GATGAGCCTGGCCAGAGGGCCTCACCTCCACCTCCCTGCTAGACCTCCTGCTGCCACTGGC
TTGGAGCCACTGGACTCAGAGGAGCCTAGTGAGACCATGGGCCTGGGAGCTGGGCTGGGAGCC
TCTGGCTCAGGCTTCCCCAGCGAAGAGAATGAAGAGTCTCGGATTCTGCAGCCACCACAGTAC
TTCTGGGAAGAGGAGGAAGAGCTGAATGACTCAAGTCTGGACCTGGGACCCACTGCAGATTAT
GTTTTTCTGACTTAAGTGAAGGAGGTTCCATTGAAGACACTAGCCAGGCTCAAGAGCTG
CCAAACCTCCCCTCTCCCTTGCCCAAGATGAATCTGGTTGAGCCTCCCTGGCATATGCCTCCC
AGAGAGGAGGAAGAAGAGGAAGAGGAAGAGGAGGAGAGGGAGAAGGAAGAGGTAGAGAAACAA
GAGGAGGAGGAAGAGGAGGAGCTGCTCCCTGTGAATGGATCCCAAGAAGAAGCCAAGCCTCAG
GTCCGTGACTTTTTCTCTCACCAGCAGCAGCCAGACCCAGGGGCCACCAAAAGCAGGCATGAA
GACTCCGGGGACCAGGCCTCATCAGGTGTGGAGGTGGAGAGCAGCATGGGGCCAGCTTGCTG
CTGCCTTCAGTCACCCCAACTACAGTGACTCCGGGGGACCAGGACTCCACCAGCCAAGAGGCA
GAGGCCACAGTGCTGCCAGCTGCAGGGCTTGGGGTAGAGTTCGAGGCTCCTCAGGAAGCAAGC
GAGGAAGCCACTGCAGGAGCAGCTGGTTTGTCTGGCCAGCAGCAGGAGGTGCCGGCCTTGCT
TCATTCCCTCAAACCACAGCTCCCAGTGGGGCCGAGCAGCCAGATGAAGATCCCCTTGCTCT
AGAACCTCAGCCTCTTCCCCACTGGCCCCCTGGAGACATGGAAGTACACCTTCTCTGCTACC
TTGGGACAAGAAGATCTCAACCAGCAGCTCCTAGAAGGGCAGGAGCTGAAGCTCAATCCAGG
ATACCTTGAGATTCTACGCAGGTGATCTGCAAGGACTGGAGCAATCTGGCTGGGAAAACTAC
ATCATTCTGAACATGACAGAGAACATAGACTGTGAGGTGTTCCGGCAGCACCGGGGGCCACAG
CTCCTGGCCCTGGTGGAAGAGGTGCTGCCCCGCCATGGCAGTGGCCACCATGGGGCCTGGCAC
ATCTCTCTGAGCAAGCCCAGCGAGAAGGAGCAGCACCTTCTCATGACACTGGTGGGCGAGCAG
GGGGTGGTGCCCACTCAAGATGTCCTTTCCATGCTGGGTGACATCCGCAGGAGCCTGGAGGAG
ATTGGCATCCAGAACTATTCCACAACCAGCAGCTGCCAGGCGCGGGCCAGCCAGGTGCGCAGC
GACTACGGCACGCTCTTCGTGGTGCTGGTGGTCATTGGGGCCATCTGCATCATCATCATTCGCG
CTTGGCCTGCTCTACAAGTCTGGCAGCGCCGGCTGCCCAAGCTCAAGCACGTGTGCGACGGC
GAGGAGCTGCGCTTCGTGGAGAACGGCTGCCACGACAACCCACGCTGGACGTGGCCAGCGAC
AGCCAGTCGGAGATGCAGGAGAAGCACCCAGCCTGAACGGCGGCGGGGCCCTCAACGGCCCG
GGGAGCTGGGGGGCGCTCATGGGGGGCAAGCGGGACCCCGAGGACTCGGACGTGTTTCGAGGAG
GACACGCACCTG**TGAG**CGCAGCCGAGGCGCAGGCCGAGTGGGCCGCCAGGACCAAGCGAGGTG
GACCCCGAAACGGACGGCCCGAGCCCGCACCAGCCCCGCGCCTACCCGGGGCGCCCCCGCGG
CCTGGCCCTCGGCGCGGGCTCCTTCCCGCTTCCCCGACTTCACACGGCGGGCTTCGGACCAAC
TCCCTCACTCCCGCCCGAGGGGCAGGCCTCAAAGCCCCGCTTGGGCCCCGCTTCCCGCCCCCTG
AACCCCGGCCCGCGGGCGGGCGGGCGGCTTCTGCGCCCCGGGACTCAATTAACCCGCCC
GGAGACCACGCCGGGCCAGCAAAA

160/550

FIGURE 160

MGRLLRAARLPPLLSPLLLLLLVGGAF LGACVAGSDEPGPEGLTSTSLLDLLLPTGLEPLDSEE
PSETMGLGAGLGASGSGFPSEENEESRILQPPQYFWEEEEELNDSSLDLGPTADYVFPDLTEK
AGSIEDTSQAQELPNLPSPLPKMNLVEPPWHMPPREEEEEEEEEEREKEEVEKQEEEEEEEL
LPVNGSQEEAKPQVRDFSLTSSSQTPGATKSRHEDSGDQASSGVEVESSMGPSLLLPSVTPTT
VTPGDQDSTSQA EATVLP AAGLGVEFEAPQEASEEATAGAAGLSGQHEEVPALPSFPQT TAP
SGAEHPDEDPLGSRTSASSPLAPGDMELTPSSATLGQEDLNQQ LLEGQA AE AQSRIPWDSTQV
ICKDWSNLAGKNYIILNM TENIDCEVFRQHRGPQLLALVEEVLPRHGSGHHGAWHISLSKPSE
KEQHLLMTLVGEQGVVPTQDVLSMLGDIRRSLEEIGIQNYSTTSSCQARASQVRSDYGT L FVV
LVVIGAICIIIIALG LLYNCWQRRLPKLKHVSHGEELRFVENGCHDNPTLDVASDSQSEMQEK
HPSLNGGGALNGPGSWGALMGGKRD PEDSDVFEEDTHL

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 499-521

N-glycosylation sites.

amino acids 106-110, 193-197, 395-399, 480-484

Glycosaminoglycan attachment site.

amino acids 77-81

N-myristoylation sites.amino acids 24-30, 28-34, 41-47, 69-75, 71-77, 73-79, 75-81,
216-222, 327-333, 455-461, 519-525, 574-580, 581-587, 584-590**Amidation site.**

amino acids 588-592

161/550

FIGURE 161

CCAGGGCGGAGCGCAGCTGCGCCGGGCTTGGGCGCCTGGGGCCGCGCTCCCCACCGTCGTTT
TCCCCACCGAGGCCGAGGCGTCCCGGAGTC**ATG**GCCGGCCTGAACTGCGGGGTCTCTATCGCA
CTGCTAGGGGTCTGCTGCTGGGTGCGGCGCGCCTGCCGCGCGGGGCAGAAGCTTTTGAGATT
GCTCTGCCACGAGAAAGCAACATTACAGTTCTCATAAAGCTGGGGACCCGACTCTGCTGGCA
AAACCCTGTTACATCGTCATTTCTAAAAGACATATAACCATGTTGTCCATCAAGTCTGGAGAA
AGAATAGTCTTTACCTTTAGCTGCCAGAGTCCTGAGAATCACTTTGTCATAGAGATCCAGAAA
AATATTGACTGTATGTCAGGCCCATGTCCTTTTGGGGAGGTTCAGCTTCAGCCCTCGACATCG
TTGTTGCCTACCCTCAACAGAACTTTCATCTGGGATGTCAAAGCTCATAAGAGCATCGGTTTA
GAGCTGCAGTTTTCCATCCCTCGCCTGAGGCAGATCGGTCCGGGTGAGAGCTGCCAGACGGA
GTCACTCACTCCATCAGCGGCCGAATCGATGCCACCGTGGTCAGGATCGGAACCTTCTGCAGC
AATGGCACTGTGTCCCGGATCAAGATGCAAGAAGGAGTGAAAATGGCCTTACACCTCCCATGG
TTCCACCCCAGAAATGTCTCCGGCTTCAGCATTGCAAACCGCTCATCTATAAAACGTCTGTGC
ATCATCGAGTCTGTGTTTGAGGGTGAAGGCTCAGCAACCCTGATGTCTGCCAACTACCCAGAA
GGCTTCCCTGAGGATGAGCTCATGACGTGGCAGTTTGTCTGTTCCCTGCACACCTGCGGGCCAGC
GTCTCCTTCCCTCAACTTCAACCTCTCCAACCTGTGAGAGGAAGGAGGAGCGGGTTGAATACTAC
ATCCCGGGCTCCACCACCAACCCCGAGGTGTTCAAGCTGGAGGACAAGCAGCCTGGGAACATG
GCGGGGAACTTCAACCTCTCTCTGCAAGGCTGTGACCAAGATGCCCCAAAGTCCAGGGATCCTC
CGGCTGCAGTTCCAAGTTTTGGTCCAACATCCACAAAATGAAAGCAGTGAG**TGAG**CCCCACTT
TCCTTTTTCTTCCCTCCTCCAGCACCTTCGTTGTTTCCTGGGTAGTCTGCCTGGGTGAGGCTCC
CTTCCCTGTTTCTCATCTGTGGCTTCTGAAACACTTAGACTCTGGACCCAGCAAGAGTTTCAGG
AAGTGGGTGCTAGGCAGTTAGACAGGCTTGTTGGTGAACACCCGGTATGTAGTTCCATTTCA
GCACAATAAAAAGAAATCTTGCATTCAAGATGCTAAATTGTTTTTAACGAAAA

162/550

FIGURE 162

MAGLNCGVSIALLGVLLLGAARLPRGAFAFEIALPRESNITVLIKLGTPDLLAKPCYIVISKR
HITMLSIKSGERIVFTFSCQSPENHFVIEIQKNIDCMMSGPCPFGEVQLQPSTSLLPDLNRFTI
WDVKAHKSIGLELQFSIPRLRQIGPGESCPDGVTHSISGRIDATVVRIGTFCSNGTVSRIKMQ
EGVKMALHLPWFHPRNVSGFSIANRSSIKRLCIIESVFEGEGSATLMSANYPEGFPEDELMTW
QFVVPALHLRASVSFLNFNLSNCERKEERVEYYIPGSTTNPEVFKLEDKQPGNMAGNFNLSLQG
CDQDAQSPGILRLQFQVLVQHPQNESSE

Signal peptide:

amino acids 1-29

N-glycosylation sites.amino acids 39-43, 122-126, 180-184, 205-209, 213-217, 270-274,
310-314, 339-343**Tyrosine kinase phosphorylation site.**

amino acids 276-284

N-myristoylation sites.

amino acids 3-9, 7-13, 158-164, 175-181, 191-197, 303-309

163/550

FIGURE 163

CAACCACACACCTGGGGAATTGCTGGCCTGACTTCTGACCCCTGACTCCTCATACCCCTTCCTC
CAGAGCATGACATTTGACCACCAACTGAAACCTGACCTCTGACCCCAGACCACTGGCCCTTCC
CCCGCCCTGTGGTGACTTCATAAAGGTTACTAGCTTCTCCCCTGGCCTTGAGACCCACACGAT
GGCCCTGCTGGCTCTGGCCAGTGCCGTCCCCTCTGCCCTGCTGGCCCTGGCTGTCTTCAGGGT
GCCCCCTGGGCCTGTCTCCTCTGCTTCACAACCTACTCTGAGCGCCTCCGCATCTGCCAGAT
GTTTGTGGGATGCGGAGCCCCAAGCTTGAAGAGTGTGAGGAGGCCTTCACGGCCGCCTTCCA
GGGCCTCTCTGACACCGAAATCAGTGAGGAGACCATCCACACTTCATCAGTGTCTGGGGAAG
GTGCAGAGGGAGGGCAGGAGAGGCCAGAGGGTCAGGCTGAGGGACAGACAGAGAGAAACAGT
CAGAGGAGAAAGGCTCAAAGACCATGAGAACAACAGAGACTTAGGGACAGAGAGACACAGACA
GGGAAGACAGCAGGGCAAAGACTCAGAGAGGGGAGGATGGAGAGTCAGAGAGGGGAAGATGG
AGACTCAGAGAGAGGGGAGGATGGAGACTCAGAGAGAGAGGAAGATGGAGACTCAGAGGGAAA
GATGGAGACTCAGGAGTATGGAGAGTCAGAGAGGGGAGGATGGACACTCAGGGGAGGATGGAG
AGTCAGGAGGATGGAGACTCATAGAAAGGGGAGGATGGAGAGTCAGGAGAGGTTGGAGACTGG
AGAGGGAATAGAGACCCAGAAAGGGGAGGATGGAGACTCAGAGGGTGGAAGATGGAGACTCAA
AGAGGATGGAAACCCAGAGAGAGGAGGACAGAGATGAGGCAGAGACTAGGGGAAGCAGGATAG
CGACTGGTTCGGGGCAGAGACTCAGGGAGGATAGAGACTCACAGAGAGGTGAGGATAGAGACT
TGGGAGGGACTCAGGAAGCATAGCGACTGTGGGGCAAAGAGTCAGAGAGGGGAGGATACAGAC
TTGGGAGGGCAGAGACTCAGAAACAGAATGTTTCGATTAGGGACATGGTGTTGCGGGGAGCTG
CCTCCCCCAGCCCCTGCTCCCTCCCTCACC GCCAGACTATGATGAGAGAAGCCACCTGCATGA
CACCTTCACCCAGATGACCCATGCCCTGCAGGAGCTGGCTGCTGCCCAGGGATCCTTTGAGGT
TGCCTTCCCTGATGCTGCAGAGAAAATGAAGAAGGTCATTACACAGCTTAAAGAAGCCCAGGC
TTGCATCCCTCCCTGCGGTCTCCAGGAGTTCGCCCCGCGTTTCCTCTGCAGCGGGTGCTACTC
TAGGGTCTGCGACCTCCCGCTGGACTGCCCAGTT CAGGATGTGACAGTGACTCGGGGCGACCA
GGCTATGTTTTCTTGATCGTAAACTTCCAGCTGCCAAAGGAGGAGATCACCTATTCCTGGAA
GTTTCGAGGAGGAGGTCTCCGGACTCAGGACTTGTCCTATTTCCGAGATATGCCGCGGGCCGA
AGGATACCTGGCGCGGATCCGGCCGGCTCAGCTCACGCACCGCGGGACGTTCTCCTGCGTGAT
CAAGCAAGACCAGCGCCCCCTGGCCCCGCTCTACTTCTTTCTTAACGTCTTCGGGGCCCTCGC
ATCAGCGAGTGCGACAGTGTTGGCGTGAGTTCTGGGGACTCCGGAGCCCCAGCATCTAGC
TCCCCGCTGTCTCAGATCCCACCGAGAAGTCTGGGTTCCAGCAACCTCCAACCCAGGAGGAT
GTTCTTTTCGATGGTACTGCAGTGGCAACTAACAAAGGTATCTTTCCTCCTTCCCTATCTATT
TCCATCCTGAAAATAAAGAATATATTTCAACTCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAA

164/550

FIGURE 164

MALLALASAVPSALLALAVFRVPAWACLLCFTTYSERLRICQMFVGMRS PKLEECEEAF TAAF
QGLSDTEISEETIHTSSVSWGRRCRGRAGEAQRVRLRDRQRETVRGERLKDHENNRDLGTERHR
QGKTAGQRLREGRMESQRGEDGDSEGEDGDSEEREEDGDSEGKMETQEYGESERGGWTLRGGW
RVRRMETHRKGRMESQERLETGEGIETQKGEDGDSEGGRWRLKEDGNPERGGQR

Signal peptide:

amino acids 1-26

N-myristoylation site.

amino acids 65-71

FIGURE 165

[illegible]

166/550

FIGURE 166

MELSDVTLIEGVGNEVMVVAGVVVLILALVLAWLSTYVADSGSNQLLGAIVSAGDTSVLHLGH
VDHLVAGQGNPEPTELPHPSEGNDKAEKEAGEGRGDSTGEAGAGGGVEPSLEHLLDIQGLPKR
QAGAGSSSPEAPLRSEDSTCLPPSPGLITVRLKFLNDTEELAVARPEDTVGALKSKYFPGQES
QMKLIYQGRLLQDPARTLRSLNITDNCVIHCHRSPPGSAVPGPSASLAPSATEPPSLGVNVGS
LMVPVFVLLGVVWYFRINRQFFTAPATVSLVGVTVFFSFLVFGMYGR

Signal peptide:

amino acids 1-36

Transmembrane domains:

amino acids 246-267, 275-301

N-glycosylation sites.

amino acids 162-166, 211-215

N-myristoylation sites.

amino acids 48-54, 105-111, 109-115, 129-135, 177-183, 247-253

Cell attachment sequence.

amino acids 97-100

167/550

FIGURE 167

GGCGGCTGTGTGTCGCCGGAGCCGAAGCGCGCAGGCCCGTCCCGGTGGCCGGGAGCGGGCGGGTGGGGGCGCCA
TGTGGTTTCATGTACCTGCTGAGCTGGCTGTGCTCTTCATCCAGGTGGCCTTCATCACGCTGGCTGTGCGGGCTG
GACTCTATTACCTGGCAGAACTGATAGAAGAATACACAGTGGCCACCAGCAGGATCATAAAATACATGATCTGGT
TCTCCACCGTGTACTGATTGGCCTCTACGTCTTTGAGCGCTTCCCCACCAGCATGATTGGAGTGGGCCTATTCA
CCAACCTCGTCTACTTTGGCCTCCTCCAGACCTTCCCCTTCATCATGCTGACCTCGCCTAACTTCATCCTGTGCT
GTGGACTAGTGGTGGTGAATCATTACCTAGCATTTCAGTTTTTGCAGAAGAATATTATCCCTTCTCAGAGGTCC
TGGCCTATTTCACCTTTCTGCCTGTGGATAATTCCGTTTGCCTTTTTTGTGTCACTTTTCGGCCGGGAGAACGTCC
TGCCCTCTACCATGCAGCCAGGAGATGATGTCGTCTCCAATTATTTACCAAAGGCAAGCGGGGCAAACGCTTAG
GGATCCTGGTGTCTTCTCCTTCATCAAAGAGGCCATTTACCCAGTTCGTGAGAAGATATACTGACCCCCATGCA
GGCAGGATGTGGGGGGCAAGATCAGGAGAGTCAGGCCCTGGGCCCTCATGCCAGGTGGGGACCAGAAGTCGGGA
AGGCACCTACCACCTGCCCTGGCTTTCTTCCCCTCAACTCTGGAGCCCCATCCCCACCCTCCTTGGGGGGCTCAG
CTTGGCTCAGATCTGATGCTTCAAGAGGCTGTAACCTCAGAGGGCACCAGGAGGGTGGCAGAGCCTGCTTAGCC
AGGAGGCCGAGGTCCCTCAGTCCTCCCCTGTCCCCTCCAAGGTGGGTGTCAGGAGGTCTGGCCCCGCTGGGGCAGG
CAGGGCAGGCTCTGTGAAGCTTAAGAGCAGATGGTGACAAGTTCTCTGGGCAGGTGGCCATGGGGAGGGGCATG
GCTTGGCATGTCCAACAGAAATAGTTTTTGGCTGTTGAACGGTGATTCTGTCCAAGTGCAGATTTCCGTTTGAAT
AAAGCTTCGCTCTAGGTGGCACTGTTTGCCTTAATACCCTGACAGTTCATCTTCCCTTCTCCTGCTAACCTTC
TGCTCTGGACTGGACTCACTTTTCTGCTCCAGGACTCCTTTTCTGGGTTTGGGTCTTGCCCTTCCCAAGGGACT
GTTCTTGTGGCCCTTAATGGGAAGGGGGCAGGGGTGAGGAGCTGAGCCTGCTCAAGGAGTGGGAAGTGGGGCTAT
AGGCAGCCTCTCTGATGCACCTCTTCCATCTCTTCCCCAAGGCTCCGTGACTGTCAAACCTGGGAGTAGGAGAG
GGGACAATTTAGGACTGGGCTAGATTTTTCAGAAAGAACATCTACAATATCCTATTTATAAATCTTCTCTGGGAAA
AGGAGTGGTTTCTGGCTGAATACTATCTTAGGCTCAAGGAGAAACAAAATAAAAATTAGCTTCCAGGCAGCCTGT
TTTTAAAGAAATGGGACTAATGGGAGAAGCTGTTTGTCACTCTAAGAGCATCCAAGCCCTGGCCCCGTCTGTGCAC
TCTTGGCTCCTGGGGAGATATATCTGCCTTCTAAGAAGGCAGGCCAGGTCTTGGGCACAGACCTGCATTTGTTGA
CCTTGCACTCCAATGATAGTGCCTTGCAAGTGCTCAACAGTACATATTGGAATGAAGTCCCTATGAGAGCCATTT
CTGGCCATGTTCTATACCTCAAAGTGAGGCTGGCAGGTACAGAGATGAACTGTACACATGTGATACATTTAAGCC
ACTGGAAAAACCCCTGTGCTTGAATAATTTCTCTATATCATGCCTGGAGTTCCATCATAGCCCTTCATTTCT
TGGCTTTAGCATTTACCTTCTCTTAAGAATACCAGCTTTCCCCTTTCCCTGAGAGGAAGAGCACATGTTGGTCTC
CTCTTAGTGTGAACGAGATTGCCAGGCCCTTTTCTCCTATGCACACCAGGATAGACAAGGCAGGGGATACTGGCA
GCCTGCATCATCCTCCCATTGGGCTGACAGCTGGCCCTACTTTCCCTCCCTCTGCTGCTTGGTCCCTCACCTTGAT
GATGTGGCTTCGCCCCCTCCACTCTACTGCCAGTGTTCTCCAGGGGTTGCTAAATCCAGCAGACCCCTTTCTCTG
TCTTACTAGATCTGGGCAGCATTTGACATGGCTGATCACCCCTTGCTTCTTGGATGGCACTTCCCTGGCACCTCT
GTGGCTAGTTGTCTACCTCCCTGGCTGTTCCCTTCAGGCTTCCGTGCAGGCTTCTCCACTTGCCCATGCACAGT
AGGGTCTTTAGGGTCTGTGCTGTGGGCTCCCTAGGGAAGCCCATCCATCTGGATGGTTTCAAGGATGGTGAGGAA
TTTAGAGTTGACCTCCAGCCCCAACATCCTTCTGATCACTGAACCACAGTTTTGCTGCCCTCTAGGTGCACAG
ACAATTCAGGTCCATGGCCAGATGGTACTTGCTGTCTTCTGCAAACCTGCCCTTCTGGGTACTTCCCTTGACC
CCGAGATCACTCAGGAGCCAGACAGGAACTTATTCTATTCTGTTTTCTCTTTCTGCCACCACATCCAATCTC
TCAAACCGTCAGGTCTACCTTAACATCTCTTGATTTGAGCCACTCCCACTGTCATCAGCTTTCACCTGGATTAT
CGTGACAGCCTCCTACTGCTTCTCTATCATGTGGCCAGAGCTATCTTCTAAAATGCATTGCATAGTTGATCAAG
TCACTCTCTGGCCTAAAACCTTCCCTGGCTCCCTGCTGCCCTCAGGATAAAGTCTGGACCCCTCAGCATGGCTTG
TGAGACTCATGGTGTCTTGTCCCTGCTCACCTCTCTGGTCTCATCACTTGCCTTCTTGCATTCTGGGTCCCAGC
CTCCTGTATCCAGAGATGCAGTGGCTCTCATTGCCACTCTGATTCCTCCTTTCTTTTGGTTCACAGAGAAAGGGT
ACTTTCTCTGTCAAATCTCAACTTAGACTTGACTTCCCTCCAAGGAGCTTTGGCTATACTCTCTCTCCCGACCCC
CACCCTGGCATACTACACAGATCACTCTGGGCTCACTTGCCTGCCTAATGGTCATCTCCCAGTAGACTGTAAGC
TCCTTGAGGGCAAGGATTGTGTTGGAATTTTTGTATTAAACAGTGCCCTGGCTTGGTGCCTGGCACCTAGAAAGCAC
TCAATAAATGTTTGTTTAATGAA

168/550

FIGURE 168

MWFMYLLSWLSLFIQVAFITLAVAAGLYYLAELIEEYTVATSRIIKYMIWFSTAVLIGLYVFE
RFPTSMIGVGLFTNLVYFGLLQTFPFIMLTSPNFILSCGLVVVNHYLAFQFFAEEYYPFSEVL
AYFTFCLWIIIPFAFFVSLSAGENVLPSTMQPGDDVVSNYFTKGKRGKRLGILVVVFSFIKEAIL
PSRQKIY

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 126-146

Casein kinase II phosphorylation site.

amino acids 145-148

N-myristoylation sites.

amino acids 73-78, 82-87

Amidation sites.

amino acids 168-171, 171-174

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 91-101

169/550

FIGURE 169

CAAAGCCCTACCCTCACCATTACCCAGGTCCTGTGGGAAGAGCAGCGTGGAGGTGGGCTGAGG
TTAGAAGGTGCAGAGCGTGGAAGAAGATTGTGAGCTGAGTATTGGACATCTGTTCTTGAATAG
TCCCTGGGCCTGCCATAGGAAAGGAAGTTCTCCAGGGTTACAGTTCTTATCCGCGTGAATACA
CATGGCTCTGTTACGAAAAATTAATCAGGTGCTGCTGTTCTTCTGATCGTGACCCTCTGTGT
GATTCTGTATAAGAAAGTTCATAAGGGGACTGTGCCCAAGAATGACGCAGATGATGAATCCGA
GACTCCTGAAGAACTGGAAGAAGAGATTCTGTGGTGATTTGTGCTGCAGCAGGGAGGATGGG
TGCCACTATGGCTGCCATCAATAGCATCTACAGCAACACTGACGCCAACATCTTGTTCTATGT
AGTGGGACTCCGGAATACTCTGACTCGAATACGAAAATGGATTGAACATTCCAACTGAGAGA
AATAAACTTTAAAATCGTGGAATTCAACCCGATGGTCCTCAAAGGGAAGATCAGACCAGACTC
ATCGAGGCCTGAATTGCTCCAGCCTCTGAACTTTGTTGATTTTATCTCCCTCTACTTATCCA
CCAACACGAGAAAGTCATCTATTTGGACGATGATGTAATTGTACAAGGTGATATCCAAGAACT
GTATGACACCACCTTGCCCTGGGCCACGCGGCGGCTTTCTCAGATGACTGCGATTTGCCCTC
TGCTCAGGACATAAACAGACTCGTGGGACTTCAGAACACATATATGGGCTATCTGGACTACCG
GAAGAAGGCCATCAAGGACCTTGGCATCAGCCCCAGCACCTGCTCTTTCAATCCTGGTGTGAT
TGTTGCCAACATGACAGAATGGAAGCACCAGCGCATCACCAGCAATTGGAGAAATGGATGCA
AAAGAATGTGGAGGAAAACCTCTATAGCAGCTCCCTGGGAGGAGGGGTGGCCACCTCCCCAAT
GCTGATTGTGTTTCATGGGAAATATTCCACAATTAACCCCTGTGGCACATAAGGCACCTGGG
CTGGAATCCAGATGCCAGATATTCGGAGCATTTTCTGCAGGAAGCTAAATTACTCCACTGGAA
TGGAAGACATAAACCTTGGGACTTCCCTAGTGTTTACAACGACTTATGGGAAAGCTGGTTTGT
TCCTGACCCTGCAGGGATATTTAACTCAATCACCATAGCT**TGAT**AATACTCTACCCTTAAAAT
ATTCCCTGTATAGAAATGTGGAATTGTCCCTTTGTAGCCAACATAACATTGTTCTTTATGAA
TATTACCTTTGATACATATGATCCACAATATAAAAACCAAAAACACTACTGTGTGCAAATTATAC
CTTGGACCATATAGGCATTGATTAACCTCTTTAAGTACATGTGATAACTATGGAAATCAAGAT
TATGTGACTGAAAAACATAAAGGAAGAGACCCATCTAGATAACAGCAATCAACCTGCTTAATT
CTGAATGACAATTATATCCACAAATTTTTAAAACCTCTACATGTATTTTTCATGAAGATCT
CCTTAACAGGTTGCCAACCTTTCTTTTATAAACTATTACATTTAAAATATGGACGTCTGAA
AAATAAAATATTCATCATTTTTTAAAA

170/550

FIGURE 170

MALLRKINQVLLFLLIVTLCVILYKKVHKGTVPKNDADDESETPEELEEIPVVICAAAGRMG
ATMAAINSIYSNTDANILFYVVGLRNTLTRIRKWIEHSLREINFKIVEFNPMVLKGKIRPDS
SRPELLQPLNFVRFYLPLLIHQHEKVIYLDLDDVIVQGDIQELYDTTLALGHAAAFSDDCDLPS
AQDINRLVGLQNTYMGYLDYRKKAIKDLGISPSTCSFNPGVIVANMTEWKHQIRITKQLEKWMQ
KNVEENLYSSSLGGGVATSPMLIVFHGKYSTINPLWHIRHLGWNPDARYSEHFLQEAKLLHWN
GRHKPWDFPSVHNDLWESWFVPDPAGIFKLNHHS

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 234-238

Tyrosine kinase phosphorylation site.

amino acids 253-261

N-myristoylation sites.

amino acids 63-69, 86-92, 198-204, 218-224, 229-235, 265-271,
266-272

171/550

FIGURE 171

GCCAGAGGCTGCAGCTGGAGCCCAGAGCCCAAG**ATG**GAGCCCCAGCTGGGGCCTGAGGCTGCC
GCCCTCCGCCCTGGCTGGCTGGCCCTGCTGCTGTGGGTCTCAGCCCTGAGCTGTTCTTTCTCC
TTGCCAGCTTCTTCCCTTTCTTCTCTGGTGCCCCAAGTCAGAACCAGCTACAATTTTGGAAGG
ACTTTCCTCGGTCTTGATAAATGCAATGCCTGCATCGGGACATCTATTTGCAAGAAGTTCTTT
AAAGAAGAAATAAGATCTGACAACCTGGCTGGCTTCCCACCTTGGACTGCCTCCCGATTCCCTTG
CTTTCTTATCCTGCAAATTACTCAGATGATTCCAAAATCTGGCGCCCTGTGGAGATCTTTAGA
CTGGTCAGCAAATATCAAAACGAGATCTCAGACAGGAGAATCTGTGCCTCTGCATCAGCCCCA
AAGACCTGCAGCATTGAGCGTGTCTGCGGAAAACAGAGAGGTTCCAGAAATGGCTGCAGGCC
AAGCGCCTCACGCCGGACCTGGTGCAGGACTGTCACCAGGGCCAGAGAGAACTAAAGTTCCTG
TGTATGCTGAGAT**TAA**CACCAGTGAAAAAGCCTGGCATGGAGCCCAGCACTGAGAACTTCCAGA
AAGTGTTAGCCTTCTCCCAACTGTGTTATACCAACCACATTTTCAAATAGTAATCATTAAAGA
GGCTTCTGCATCAA

172/550

FIGURE 172

MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSSLVPQVRTSYNFGRTFLGLDKCNAC
IGTSICKKFFKEEIRSDNWLASHLGLPPDSLLSYPANYSDDSKIWRPVEIFRLVSKYQNEISD
RRICASASAPKTCSIERVLRKTERFQWLQAKRLTPDLVQDCHQGQRELKFLCMLR

Signal peptide:

amino acids 1-28

N-glycosylation site.

amino acids 100-103

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 158-161

N-myristoylation sites.

amino acids 56-61, 65-70

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 18-28

Prenyl group binding site (CAAX box).

amino acids 179-182

Leucine zipper pattern.

amino acids 5-26

173/550

FIGURE 173

GCTGGACTGCTCGCTGGCCGGCAGCGCACCGTTTTGAAGGTCCTAGCCCACCTGGGCTGGCTC
ACGCGCACGACTAGCCGCTCCCATACAGCACGCCCGGACTCTGTCGTCGCTTAAGGCCACTCC
TATTCTACGGCTGACCCCTGGTGGTCACGTGGATCTGTTCCGCCACGCAAGTCTGGGTCCTTCG
GCGATTGACCGGGGTCCTTGCTGTTCTGGGAGCCTCTCCTAAGCTGCCTGTTCTGCGCGAGAGTT
TGGAGGGGCGGGTTTGGGGTCGGTGTCTGATTGGGGCTCGCACCGCAGCACGCTGGAGTCCCCG
CTTAGGTACCAAGTTAGCGTCAGGGGAGCTGGGTCAGGCGGTCGCCGGGACACCCCGTGTGTGG
CAGGCGGCGAAGCGCTCTGGAGAATCCCGGACAGCCCTGCTCCCTGCAGCCAGGTGTAGTTTC
GGGAGCCACTGGGGCCAAAGTGAGAGTCCAGCGGTCTTCCAGCGCTTGGGCCACGGCGGGCGGC
CCTGGGAGCAGAGGTGGAGCGACCCATTACGCTAAAGATGAAAGGCTGGGGTTGGCTGGCCC
TGCTTCTGGGGGCCCTGCTGGGAACCGCCTGGGCTCGGAGGAGCCAGGATCTCCACTGTGGAG
CATGCAGGGCTCTGGTGGATGAACTAGAATGGGAAATTGCCCAGGTGGACCCCAAGAAGACCA
TTCAGATGGGATCTTTCCGGATCAATCCAGATGGCAGCCAGTCAGTGGTGGAGGTGCCTTATG
CCCCTCAGAGGCCCCACCTCACAGAGCTGCTGGAGGAGATATGTGACCGGATGAAGGAGTATG
GGGAACAGATTGATCCTTCCACCCATCGCAAGAACTACGTACGTGTAGTGGGCCGGAATGGAG
AATCCAGTGAACTGGACCTACAAGGCATCCGAATCGACTCAGATATTAGCGGCACCCTCAAGT
TTGCGTGTGAGAGCATTGTGGAGGAATACGAGGATGAACTCATTGAATTCTTTTCCCGAGAGG
CTGACAATGTTAAAGACAACTTTGCAGTAAGCGAACAGATCTTTGTGACCATGCCCTGCACA
TATCGCATGATGAGCTATGAACCACTGGAGCAGCCCACACTGGCTTGATGGATCACCCCCAGG
AGGGGAAAATGGTGGCAATGCCTTTTATATATTATGTTTTTACTGAAATTAAGTGAATAATA
TGAAACCAAAAGT

174/550

FIGURE 174

MKGWGWLALLLGALLGTAWARRSQDLHCGACRALVDELEWEIAQVDPKKTIQMGSFRINPDGS
QSVVEVPYARSEAHLTELEEICDRMKEYGEQIDPSTHRKNYVRVVGRNGESSELDLQGIRID
SDISGTLKFACESIVEEYEDELIEFFSREADNVKDKLCSKRTDLCDHALHISHDEL

Signal peptide:

amino acids 1-20

N-myristoylation sites.

amino acids 12-18, 16-22, 29-35

Endoplasmic reticulum targeting sequence.

amino acids 179-184

175/550

FIGURE 175

CGCAGCGCGGCAGTCCTGATGCCCCGGCATGGGTTACCGCTGCTGCCCCTGCTGTGCTCCTG
GTCGGCGCGTGGCTCAAGCTAGGAAATGGACAGGCTACTAGCATGGTCCAAGTGCAGGGTGGG
AGATTCTGATGGGAACAAATTCTCCAGACAGCAGAGATGGTGAAGGGCCTGTGCGGGAGGCG
ACAGTGAAACCCTTTGCCATCGACATATTTCTGTACCAACAAAGATTTTCAGGGATTTTGTG
AGGGAGAAAAAGTATCGGACAGAAGCTGAGATGTTTGGATGGAGCTTTGTCTTTGAGGACTTT
GTCTCTGATGAGCTGAGAAACAAAGCCACCCAGCCAATGAAGTCTGTACTCTGGTGGCTTCCA
GTGGAAAAGGCATTTTGGAGGCAGCCTGCAGGTCTGGCTCTGGCATCCGAGAGAGACTGGAG
CACCCAGTGTTACACGTGAGCTGGAATGACGCCCGTGCCTACTGTGCTTGGCGGGGAAAACGA
CTGCCCACGGAGGAAGAGTGGGAGTTTGCCGCCCGAGGGGGCTTGAAGGGTCAAGTTTACCCA
TGGGGGAACTGGTTCCAGCCAAACCGCACCAACCTGTGGCAGGGAAAGTTCCCCAAGGGAGAC
AAAGCTGAGGATGGCTTCCATGGAGTCTCCCCAGTGAATGCTTTCCCCGCCGAGAACAACCTAC
GGGCTCTATGACCTCCTGGGGAACGTGTGGGAGTGGACAGCATCACCGTACCAGGCTGCTGAG
CAGGACATGCGCGTCCTCCGGGGGGCATCCTGGATCGACACAGCTGATGGCTCTGCCAATCAC
CGGGCCCGGGTCAACCACAGGATGGGCAACACTCCAGATTTCAGCCTCAGACAACCTCGGTTTC
CGCTGTGCTGCAGACGCAGGCCGGCCGCCAGGGGAGCTGTAAGCAGCCGGGTGGTGACAAGGA
GAAAAGCCTTCTAGGGTCACTGTCAATTCCCTGGCCATGTTGCAAACAGCGCAATTCCAAGCTC
GAGAGCTTCAGCCTCAGGAAAGAACTTCCCCTTCCCTGTCTCCCATCCCTCTGTGGCAGGCGC
CTCTCACCAGGGCAGGAGAGGACTCAGCCTCCTGTGTTTTGGAGAAGGGGGCCCAATGTGTGTT
GACGATGGCTGGGGGCCAGGTGTTTCTGTTAGAGGCCAAGTATTATTGACACAGGATTGCAAA
CACACAAACAGTTGGAACAGAGCACTCTGAAAGGCCATTTTTTAAAGCATTTTAAAATCTATTC
TCTCCCCCTTTCTCCCTGGATGATTTCAGGAAGCTGACATTGTTTCTCAAGGCAGAATTTTCC
TGTTTCTGTTTTCTCAGCCAGTTGCTGTGGAAGGAGAATGCTTTCTTTGTGGCCTCATCTGTG
GTTTCGTGTCCCTCTGAAGGAACTAGTTTCCACTGTGTAACAGGCAGACATGTAAGTATTTA
AAGCACAGTTTCAGTCCTAAAAGGGTCTGGGAGAACCAGATGATGTACTAGGTGAAGCATTGCA
TTGTGGGAATCACAAAGCAAATAGTACTCCAGAAAGACAAATATCAGAAGCTTCTATTCTTT
TTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGGGTCTTTCTCTGTTGCCAGGCTAGAGTGCCTG
GTGATCACGGCTCACTCTAGCCTTGAATTCCTGGGCCCAAGCAATTCTCCACCTCAGCCTCC
TGAGTAGCTGGGACTACAAGTGTGACCACCATGCCTGGCTAATTTTTTGAATTTTTGTAGTG
ATGGGATCTCGCTCTGTTGCCAGGGTGGTCTCGAACTCCTGGCCTCAAGCGATCCTCCCACC
TCGACCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCTCGCCTGGGCCCCCTTCTCCATA
TGCCTCCAAAAACATGTCCCTGGAGAGTAGCCTGCTCCCACTGTCACTGGATGTCATGGGG
CCAATAAAATCTCCTGCAATTGTGTATCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAA

176/550

FIGURE 176

MARHGLPLLPLLSLLVGAWLKLGNQATSMVQLQGGRFLMGTSNPSDSRDGEGPVREATVKPFA
IDIFPVTNKDFRDFVREKKYRTEAEMFGWSFVFEDFVSDLRNKATQPMKSVLWWLPVEKAFW
RQPAGPGSGIRERLEHPVLHVSWN DARAYCAWRGKRLPTEEEWEFAARGGLKGQVYPWGNWFQ
PNRTNLWQGKFPKGDKAEDGFHGVSPVNAFPQNNYGLYDLLGNVWEWTASPYQAAEQDMRVL
RGASWIDTADGSANHRARVTTRMGNTPDSASDNLGFRCAADAGRPPGEL

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 191-195

N-myristoylation sites.

amino acids 23-29, 25-31, 175-181

Amidation site.

amino acids 159-163

177/550

FIGURE 177

GCCTTCTCGCGCCTGACCATGCACCCCTGCATCTTCCTGCTGGGCCACAGGCGAGCGCTTTAT
TTCTGGAGCTGAGGGCTAAACTTTTTTGACTTTTCTTCTCCTCAACATCTGAATC**ATGCCAT**
GTGCCCAGAGGAGCTGGCTTGCAAACCTTTCCGTGGTGGCTCAGCTCCTTAACCTTTGGGGCGC
TTTGCTATGGGAGACAGCCTCAGCCAGGCCCGGTTTCGCTTCCCGGACAGGAGGCAAGAGCATT
TTATCAAGGGCCTGCCAGAATACCACGTGGTGGGTCCAGTCCGAGTAGATGCCAGTGGGCATT
TTTTGTCAATATGGCTTGCACTATCCCATCACGAGCAGCAGGAGGAAGAGAGATTTGGATGGCT
CAGAGGACTGGGTGTACTACAGAATTTCTCACGAGGAGAAGGACCTGTTTTTTAACTTGACGG
TCAATCAAGGATTTCTTTCCAATAGCTACATCATGGAGAAGAGATATGGGAACCTCTCCCATG
TTAAGATGATGGCTTCCTCTGCCCCCTCTGCCATCTCAGTGGCACGGTTCTACAGCAGGGCA
CCAGAGTTGGGACGGCAGCCCTCAGTGCCTGCCATGGACTGACTGGATTTTCCAACCTACCAC
ATGGAGACTTTTTTCATTGAACCCGTGAAGAAGCATCCACTGGTTGAGGGAGGGTACCACCCGC
ACATCGTTTACAGGAGGCAGAAAGTTCCAGAAACCAAGGAGCCAACCTGTGGATTAAAGGGTA
TTGTGACTCACATGTCCTCCTGGGTGAAGAATCTGTTTTGTTCTTTTGG**TAG**TTTTTATTAAA
ACATGACCTATTCTTACTCAAGTCTCTTATCTCCTCTGTATTCTTTTTTTTTTAATATCTTCA
TGACATTCAAATCTCTTCTGTATTCTCTTGCCAGAAAGTGACATTCTTTTTTGCTTGTATAAA
CCCTTTCACCTTGTC

178/550

FIGURE 178

MPCAQRSWLANLSVVAQLLNFGALCYGRQPQPGPVRFPDRRQEHFIKGLPEYHVVGPPVRVDAS
GHFLSYGLHYPTSSRRKRDLDGSEDWVYYRISHEEKDLFFNLTVNQGFLSNSYIMEKRYGNL
SHVKMMASSAPLCHLSGTVLQQGTRVGTAALSACHGLTGFFQLPHGDDFFIEPVKKHPLVEGGY
HPHIVYRRQKVPETKEPTCGLKGIVTHMSSWVEESVLFFW

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 11-15, 105-109, 125-129

N-myristoylation site.

amino acids 149-155

179/550

FIGURE 179

CAGATTTAAAAAGAAAACCTTTACTGAATCAGCTGAGTGTTAATAATACGAATTCCTTTTCT
TGCCAATTCTGATCTGAACAGAAAATCCAAGAACAGGGATATGCTGTGGATTACAGTTTTCTCT
GCCTTGCCTACGACTGTTTCTGGTTGTTACCTGTTATCTTTTATTATTACTCCACAAAGAAAT
ACTTGGATGTTCTGCTGTTTGTGAGCTCTGCACTGGGAGACAAAATTAAGTCCGTAAGTCTTAGG
CCTTTCGAGTATTCCTAAGAATTTTCCTGAAAGTACAGTTTTTCTGTATCTGACTGGGAATAA
TATATCTTATATAAATGAAAGTGAATTAACAGGACTTCATTCTCTGTAGCATTGTATTTGGA
TAATTCTAACATTCTGTATGTATATCCAAAAGCCTTTGTTCAATTGAGGCATCTATATTTTCT
ATTTCTAAATAATAATTTTCATCAAACGCTTAGATCCTGGAATATTTAAGGGACTTTTAAATCT
TCGTAATTTATATTTACAGTATAATCAGGTATCTTTTGTTCGAGAGGAGTATTTAATGATCT
AGTTTCAGTTCAGTACTTAAATCTACAAAGGAATCGCCTCACTGTCCTTGGGAGTGGTACCTT
TGTTGGTATGGTTGCTCTTCGGTACTTGATTTATCAAACAATAACATTTTGAGGATATCAGA
ATCAGGCTTTCAACATCTTGAAAACCTTGCTTGTGTTGATTTAGGAAGTAATAATTTAACAAA
AGTACCATCAAATGCCTTTGAAGTACTTAAAAGTCTTAGAAGACTTTCTTTGTCTCATAATCC
TATTGAAGCAATACAGCCCTTTGCATTTAAAGGACTTGCCAATCTGGAATACCTCCTCCTGAA
AAATTCAGAATTAGGAATGTTACTAGGGATGGGTTTAGTGGAATTAATAATCTTAAACATTT
GATCTTAAGTCATAATGATTTAGAGAATTTAAATCTGACACATTCAGTTTGTTAAAGAATTT
AATTTACCTTAAGTTAGATAGAAACAGAATAATTAGCATTGATAATGATACATTTGAAAATAT
GGGAGCATCTTTGAAGATCCTTAATCTGTCAATTAATAATCTTACAGCCTTGCATCCAAGGGT
CCTTAAGCCGTTGTCTTCATTGATTCATCTTCAGGCAAATCTAATCCTTGGGAATGTAAGT
CAAACTTTTGGGCCTTCGAGACTGGCTAGCATCTTCAGCCATTACTCTAAACATCTATTGTCA
GAATCCCCCATCCATGCGTGGCAGAGCATTACGTTATATTAACATTACAAATTGTGTTACATC
TTCAATAAATGTATCCAGAGCTTGGGCTGTTGTAAAATCTCCTCATATTCATCACAAGACTAC
TGCGCTAATGATGGCCTGGCATAAAGTAACCACAAATGGCAGTCCTCTGGAAAATACTGAGAC
TGAGAACATTACTTTCTGGGAACGAATTCCTACTTCACCTGCTGGTAGATTTTTTCAAGAGAA
TGCCCTTGGTAATCCATTAGAGACTACAGCAGTGTTACCTGTGCAAATACAACCTACTACTTC
TGTTACCTTGAACCTGGAAAAAACAGTGCTCTACCGAATGATGCTGCTTCAATGTCAGGGAA
AACATCTCTAATTTGTACACAAGAAGTTGAGAAGTTGAATGAGGCTTTTGACATTTTGCTAGC
TTTTTTCATCTTAGCTTGTGTTTTAATCATTTTTTTGATCTACAAAGTTGTTGAGTTTAAACA
AAAATAAAGGCATCAGAAAACCAAGGGAAAATAGACTTGAATACTACAGCTTTTATCAGTC
AGCAAGGTATAATGTAAGTGCCTCAATTTGTAACACTTCCCCAAATCTCTAGAAAGTCCTGG
CTTGGAGCAGATTGACTTCATAAACAAATGTTTCTGAAAATGAGGCACAGGTCAATCTTTT
TGAACATTCTGCTTTATTAAGTCAACTAAATATTGTCTATAAGAACTTCAGTGCCATGGACAT
GATTTAAAGTGAACCTCCTTATATAATTATATACTTTAGTTGGAAATATAATGAATTATATG
AGGTTAGCATTATTAAAATATGTTTTTNTTAAAAA

180/550

FIGURE 180

MCGLQFSLPCLRLFLVVTCTYLLLLLHKEILGCSSVCQLCTGRQINCRNLGLSSIPKNFPESTV
FLYLTGNNISYINESELTGLHSLVALYLDNSNILYVYPKAFVQLRHLYFLFLNNNFIKRLDPG
IFKGLLNLRNLYLQYNQVSFVPRGVFNDLVSVQYLNLRNRLTVLGSGTFVGMVALRILDLSN
NNILRISESGFQHLENLACLYLGSNNLTKVPSNAFEVLKSLRRLSLSHNPIEAIQPFQFAFKGLA
NLEYLLLKNSRIRNVTRDGFSGINNLKHLILSHNDLENLNSDTFSLLKNLIYKLDNRNRIISI
DNDTFENMGASLKILNLSFNNLTALHPRVLKPLSSLIHLQANSNPWECNCKLLGLRDWLASSA
ITLNIYCQNPPSMRGRALRYINITNCVTSSINVSRAWAVVKSPHIHHKTTALMMAWHKVTNG
SPLENTETENITFWERIPTSPAGRFFQENAFGNPLETTAVLPVQIQLTTSVTLNLEKNSALPN
DAASMSGKTSLECTQEVEKLNEAFDILLAFFILACVLIIIFLIYKVQFKQKLKASENSRENRL
EYYSFYQSARYNVTASICNTSPNSLESPGLEQIRLHKQIVPENEAQVILFEHSAL

Signal peptide:

amino acids 1-41

Transmembrane domain:

amino acids 530-547

N-glycosylation sites.amino acids 71-75, 76-80, 215-219, 266-270, 317-321, 331-335,
336-340, 400-404, 410-414, 451-455, 579-583**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 231-235

N-myristoylation sites.

amino acids 3-9, 69-75, 126-132, 174-180

ATP/GTP-binding site motif A (P-loop).

amino acids 506-514

181/550

FIGURE 181

GGCCTGGCGCGGCGCTCCGGTAAGGCGTGTGTGCGGCAGGGCGGGGACAGAACCGTCTCTCG
GGCTCTGGGCGTGTCCGAGACCGCGCTCCCCGCCGAAATCAAGCTCCGAGTCATCCGTGTGGG
GCATTGCTCCCCCTGGCACAGTTGGCCTCTTTCCAGAAGCCCGTTTTGTTTGTTTACGTCT
AAATTGCGCTCGGTTCTTATTTCTCTCCCTGGCAAGGTCTGAAGACGGGTAGGAGAATAACCT
GTGTCAGCGTGTT**ATG**ATGCCGTCCCGTACCAACCTGGCTACTGGAATCCCCAGTAGTAAAGT
GAAATATTCAAGGCTCTCCAGCACAGACGATGGCTACATTGACCTTCAGTTTAAGAAAACCCC
TCCTAAGATCCCTTATAAGGCCATCGCACTTGCCACTGTGCTGTTTTTGATTGGCGCCTTTCT
CATTATTATAGGCTCCCTCCTGCTGTCAGGCTACATCAGCAAAGGGGGGGCAGACCGGGCCGT
TCCAGTGCTGATCATTGGCATTCTGGTGTTCCTACCCGGATTTTACCACCTGCGCATCGCTTA
CTATGCATCCAAAGGCTACCGTGGTTACTCCTATGATGACATTCCAGACTTTGATGAC**TAG**CA
CCCACCCCATAGCTGAGGAGGAGTCACAGTGGAAGTGTCCCAGCTTTAAGATATCTAGCAGAA
ACTATAGCTGAGGACTAAGGAATTCTGCAGCTTGACAGATGTTTAAGAAAATAATGGCCAGATT
TTTTGGGTCCTTCCCAAAGATGTTAAGTGAACCTACAGTTAGCTAATTAGGACAAGCTCTATT
TTTCATCCCTGGGCCCTGACAAGTTTTTCCACAGGAATATGTATCATGGAAGAATAGAGGTAA
TTCTGTAATGGAAAAGTGTTCCTGCCACCACCTCTGTAGAGCTGAGCATTCTTTTAAATA
GTCTTCATTGCCAATTTGTTCTTGTAGCAAATGGAACAATGTGGTATGGCTAATTTCTTATTA
TTAAGTAGTTTATTTTAAAAATATCTGAGTATATTATCCTGTACACTTATCCCTACCTTCATG
TTCCAGTGGAAGACCTTAGTAAATCAAAGATCAGTGAGTTCATCTGTAATATTTTTTTTACT
TGCTTTCTTACTGACAGCAACCAGGAATTTTTTTATCCTGCAGAGCAAGTTTTCAAATGTAA
ATACTTCCTCTGTTTAAACAGTCCTTGGACCATTCTGATCCAGTTCACCAGTAGGTTGGACAGC
ATATAATTTGCATCATTTTTGTCCCTTGTAATCAAGATGTTCTGCAGATTATTCCTTTAACGG
CCGGACTTTTGGCTGTTTCCTAATGAAACATGTAGTGGTTATTATTTAGAGTTTATAGCCGTA
TTGCTAGCACCTTGTAGTATGTCATCATTCTGCTCATGATTCCAAGGATCAGCCTGGATGCCT
AGAGGACTAGATCACCTTAGTTTATTCTATTTTTTAGCTTGCAAAAAGTGACTTATATTCCA
AAGAAATTAAAATGTTGAAATCCAAATCCTAGAAATAAAATGAGTTNNNTCCAAAAAAAAAA
AAA

182/550

FIGURE 182

MMPSRTNLATGIPSSKVKYSRLSSTDDGYIDLQFKKTPPKIPYKAIALATVLFLIGAFLIIIG
SLLLSGYISKGGADRAVPVLIIGILVFLPGFYHLRIAYYASKGYRGYSYDDIPDFDD

Transmembrane domains:

amino acids 45-66, 79-95

N-myristoylation sites.

amino acids 11-17, 75-81

183/550

FIGURE 183

CTAAAAAATACAAAAATTAGCTGGGCGTGGTGTCTGTACCTGTAATCCCAGCTACTCAAGAGGCTGAGGCAGGA
GAATCGCTTGAACCCAGGAGGCAGAGGTTGCACTGAGCCAAGATTAAGTCACTGCACTCCAGCCTGGGTGACAGA
GCAAGACTCTGTATCAAAATAAATAAATAAAGTACAACCTCTGGATGGGCATGGTGGCTTATGTCTGTAATCCCAG
CACTTTGGGAACCTTGAGGCGGGTAGATTGCTTGAGTCGGGAGTTTGAGACCAGTCTGGGTAATATGGTAACCCCT
GTCTACCAAAAAATACAGGTATTAGCCAGTCTCATAACTCGGTCTCAAAATAAATAAATACATACATACATAGATG
AAAATTTAAAAATAAAGTCCAACCTCAGCGGTTTTCAGCATATTTACAGAGTTGTACAATCTTCACCACTATCTA
ATTTTCAGAACATTTTTCATCACCCCCAAAAGAAACCTAACCATTGACTATCTCTCCATTTCCCTCCTCCTCCTAG
CCTCTGGCAACCACTAATCTCTTTTTTGTCTCTATAGATTTGCCTATTTTGGACAGTTTCATATACAAGGAATCAT
ACCACATGTAGCCTTTTGTGTCCGGCTTCTTTGATTAATAGAATGTTTTCAAGGCTCATCTATGCTGTAGCCTGT
ATCAGCACTTCATTCTTTCTATGGCTGAATAATAGTCCACTGTAGGGATGTGCCATGTTTTTCCACTAGCTGAT
GGACATTTGGGTTGTTTCCACCTTCTGGCTATTATAAATATTGCTGCTATAAATATTCACTTACAAGTTTTTGTG
TGGACATATGTTTTTATTTCTTCTGGTATATCCTTCGGAGTGGAACGCTGGATCAGGTGGTAACCTCTAGGTCTA
ACCTGGCAGTTAAACAGAATCCTATGCATGCTGTAGTCCATGAGTTGAAATAAACACTTGACCCATAGTAAGTGC
CAGATCATCTTCATTTACAGCAACAGTAATTTACAGATGAGGAAATGAAGGCTCCAGAGGTGAACCTGGCTT
TTCCCATTTGAGCAGTTTCCAAGTCAGACAGTTAAAAAGTGGCAGGACCTGGAAGAGAAGCTAGTTCTTTCCACCT
GGCATTACAGGGCTGCCTCCTGGGCTACGGGGCTGGCATTTAGAATAGAGCTAAGGTCTGCTGCCAAGGCAGGTGC
CCCAGTCTGCCTCCTCTGTGCTTATTCCACTTTCTCTGCAGCCCTCCAGGGGACCCCTCTCTCAGCCACCCCTC
TCTCTGGTGAATGTCACAGTGCTGCCGGAAGATCAAAGATACGGTGCCAGAACTGGCTTCGGACCATAAGGACATT
CACAGCAGTGTATCCCGAGTGGGCAAAGCCATTGACAGGAACCTCGACTCTGAGATCTGTGGTGTGTGTGTCAGAT
GCGGTGTGGGACGCGCGGGAACAGCAGCAGCAGATCCTGCAGATGGCCATCGTGGAACACCTGTATCAGCAGGGC
ATGCTCAGCGTGGCCGAGGAGCTGTGCCAGGAATCAACGCTGAATGTGGACTTGGATTTCAAGCAGCCTTTCCTA
GAGTTGAATCGAATCCTGGAAGCCCTGCACGAACAAGACCTGGGTCTGCGTTGGAATGGGCGCTCTCCACAGG
CAGCGCTGCTGGAACCTCAACAGCTCCCTGGAGTTCAAGCTGCACCGACTGCACTTCATCCGCCTCTTGGCAGGA
GGCCCCGGAAGCAGCTGGAGGCCCTCAGTATGCTCGGCACCTTCAGCCCTTTGCTCGGCTGCACCAGCGGGAG
ATCCAGGTGATGATGGGCAGCCTGGTGTACCTGCGGCTGGGCTTGAGAGAAGTCACCCTACTGCCACCTGCTGGAC
AGCAGCCACTGGGCAGAGATCTGTGAGACCTTTACCCGGGACGCTGTTCCCTGCTGGGGCTTTCTGTGGAGTCC
CCCCTTAGCGTCAGCTTTGCCTCTGGCTGTGTGGCGCTGCCTGTGTTGATGAACATCAAGGCTGTGATTGAGCAG
CGGCAGTGCCTGGGGTCTGGAATCACAAGGACGAGTTACCGATTGAGATTGAACTAGGCATGAAGTGTGGTAC
GCTCATCTGTGGCCATGTTATCTCCCGAGATGCACTCAATAAGCTCATTAAATGGAGGAAACACTCCGTGTTCGCT
TGCCCCATCCTCCGCCAGCAGACGTGAGATTCCAACCCTCCCATCAAGCTGAAGTGTCCCTACTGTCCCATGGAG
CAGAACCCGGCAGATGGGAAACGCATCATATTCTGATTTCCTACCTGGAAGGAATTTTGTGAAAGGGGTTTTTAC
CTGTGAGCCTTGGTCTGTCTCGGTAGGTTGGTCAACTTCAGTGGACTGTGGTTGGTTTCAGAGCGCCTGGCTGAG
GAGTTCCACTGAGGGGAGCACTGGAGCAGCCCTTTGGCAGAGGCTGAGGAGGGAGATGGACCAGCCCACGCCTGG
CACCTGGCTCCATGGCATAAGGAAAGGGAGATGCTGGCCTCTGTGCTCCTGCTGTCTTTTCTGTTTCTGTTTGC
GTTTGACTTAGTAGCAACCGACAGAGTGGCAAGGGATTTGGTCTTCAGCAGTAGACATCCTTCCACCCTGCCCC
CAGCCAAGTCTCTTGTGCCATGCCAATGCTATGTCCACCCTTGCCCCCTCGGCCCCAAGAGTGTCCAGCGGTGGCC
CACCTCTTCTCCCTACAGCCTCAACAGTATGTACCATCTCCCACTGTAATAGTCCCAGTTAGAACGGAATG
CCGTTGTTTTATACTTTGAACAAATGTATTTACTGCCCTTCTCAAAA

184/550

FIGURE 184

QCCRKIKD TVQKLASDHKDIHSSVSRVGKAIDRNF DSEICGVVSDAVWDAREQQQQILQMAIV
EHLYQQGMLSVAEELCQESTLNVDLDFKQPFLELNRILEALHEQDLGPALEWAVSHRQRLLLEL
NSSLEFKLHRLHFIRLLAGGPAKQLEALSYARHFQPFARLHQREIQVMMGSLVYLRLGLEKSP
YCHLLDSSHWAEICETFTRDACSLGLSVESPLSVSFASGCVALPVL MNIAKAVIEQRQCTGVW
N HKDELPIEIELGMKCWYHSVFACFILRQQTSDSNPPIKLCGHVISRDALNKLINGGKLKCP
YCPMEQN PADGKRIIF

Transmembrane domain:

amino acids 222-241

N-glycosylation site.

amino acids 129-133

Tyrosine kinase phosphorylation site.

amino acids 151-159, 184-193

Amidation site.

amino acids 327-331

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 222-233

185/550

FIGURE 185

GAGCGACGCTGTCTCTAGTCGCTGATCCCAAATGCACCGGCTCATCTTTGTCTACACTCTAAT
CTGCGCAAACCTTTTGCAGCTGTCGGGACACTTCTGCAACCCCGCAGAGCGCATCCATCAAAGC
TTTGCGCAACGCCAACCTCAGGCGAGATGACTTGTACCGAAGAGATGAGACCATCCAGGTGAA
AGGAAACGGCTACGTGCAGAGTCCTAGATTCCCGAACAGCTACCCAGGAACCTGCTCCTGAC
ATGGCGGCTTCACTCTCAGGAGAATACACGGATACAGCTAGTGTTTGACAATCAGTTTGGATT
AGAGGAAGCAGAAAATGATATCTGTAGGTATGATTTTGTGGAAGTTGAAGATATATCCGAAAC
CAGTACCATTATTAGAGGACGATGGTGTGGACACAAGGAAGTTCCTCCAAGGATAAAATCAAG
AACGAACCAAATTAATAATCACATTCAAGTCCGATGACTACTTTGTGGCTAAACCTGGATTCAA
GATTTATTATTCTTTGCTGGAAGATTTCCAACCCGCAGCAGCTTCAGAGACCAACTGGGAATC
TGTCACAAGCTCTATTTCAGGGGTATCCTATAACTCTCCATCAGTAACGGATCCCCTCTGAT
TGCGGATGCTCTGGACAAAAAATTGCAGAATTTGATACAGTGGAAGATCTGCTCAAGTACTT
CAATCCAGAGTCATGGCAAGAAGATCTTGAGAATATGTATCTGGACACCCCTCGGTATCGAGG
CAGGTCATACCATGACCGGAAGTCAAAGTTGACCTGGATAGGCTCAATGATGATGCCAAGCG
TTACAGTTGCACTCCCAGGAATTACTCGGTCAATATAAGAGAAGAGCTGAAGTTGGCCAATGT
GGTCTTCTTTCCACGTTGCCTCCTCGTGCAGCGCTGTGGAGGAAATTGTGGCTGTGGAAGTGT
CAACTGGAGGTCTTGCACATGCAATTCAGGGAAAACCGTGAAAAAGTATCATGAGGTATTACA
GTTTGAGCCTGGCCACATCAAGAGGAGGGGTAGAGCTAAGACCATGGCTCTAGTTGACATCCA
GTTGGATCACCATGAACGATGCGATTGTATCTGCAGCTCAAGACCACCTCGATTAAGAGAATGT
GCACATCCTTACATTAAGCCTGAGAGAA

186/550

FIGURE 186

MHRLIFVYTLICANFCSCRDTSATPQSASIKALRNANLRRDDLYRRDETIQVKGNQYVQSPRF
PNSYPRNLLLTWRLHSQENTRIQLVFDNQFGLLEEAENDICRYDFVEVEDISETSTIIRGRWCG
HKEVPPRIKSRTNQIKITFKSDDYFVAKPGFKIYYSLLLEDFQPAAASETNWESVTSSISGVSY
NSPSVTDPTLIADALDKKIAEFDTVEDLLKYFNPESWQEDLENMYLDTPRYRGRSYHDRKSKV
DLDRLNDDAKRYSCTPRNYSVNIREELKLANVVFFPRCLLVQRCGNCGCGTVNWRSCNCSG
KTVKKYHEVLQFEPGHIKRRGRAKTALVDIQLDHHERCDCICSSRPPR

Signal peptide:

amino acids 1-18

N-glycosylation site.

amino acids 270-274

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 262-266

Tyrosine kinase phosphorylation site.

amino acids 256-265

N-myristoylation sites.

amino acids 94-100, 186-192, 297-303, 298-304

TonB-dependent receptor proteins signature 1.

amino acids 1-56

187/550

FIGURE 187

CATGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCTTG
TGTCTCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGT
CCTACAATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTT
CATATATGGGCAAAAGAAATGGCTGAATAAATCAGAAATGCAGAAATATCAATAGAACCTACTG
TGATCTTTCTGCTGAAACTTCTGACTACGAACACCAGTATTATGCCAAAGTTAAGGCCATTTG
GGGAACAAAGTGTTCCAAATGGGCTGAAAGTGGACGGTTCTATCCTTTTTTAGAAACACAAAT
TGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCC
AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAATATACTCCAATCT
GAAGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAA
CCACACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTACTGCGTACACGTGGAGTCCTT
CGTCCCAGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGA
TCAATCATCAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGT
GTTTCTTTTTTCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACA
CCCAGCAAATTTGATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGA
AAAAATCGTGATTAACCTTTATCACCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGA
TATGAGTTTACTGGGAAAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAA
CCTGAGGCCCCCTCAGGAGGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGA
AATTTTTTGTGACTCTGAAGAAAACACGGAAGGTACTTCTCTCACCAGCAAGAGTCCCTCAG
CAGAACAATACCCCCGGATAAAACAGTCATTGAATATGAATATGATGTCAGAACCACTGACAT
TTGTGCGGGGCCTGAAGAGCAGGAGCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATT
ATTGGAGTCGCAGGCAGCGTTGGCAGTCTTGGGCCCCGAAACGTTACAGTACTCATAACCCCC
TCAGCTCCAAGACTTAGACCCCCTGGCGCAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGA
AGAGCCATCGACGACCCTGGTCGACTGGGATCCCCAACTGGCAGGCTGTGTATTCTTCGCT
GTCCAGCTTCGACCAGGATTCAGAGGGCTGCGAGCCTTCTGAGGGGGATGGGCTCGGAGAGGA
GGGTCTTCTATCTAGACTCTATGAGGAGCCGGCTCCAGACAGGCCACCAGGAGAAAATGAAAC
CTATCTCATGCAATTCATGGAGGAATGGGGTTATATGTGCAGATGGAAAACTGATGCCAACA
CTTCCTTTTGCCTTTTGTTCCTGTGCAACAAGTGAGTCACCCCTTTGATCCCAGCCATAAA
GTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTGAGTGTCTGTGAGAA

188/550

FIGURE 188

MPLPPLLLLLLAAPWGRAVPCVSGGLPKPANITFLSINMKNVLQWTPPEGLQGKVKVTVQYF
IYGQKKWLNKSECRNINRTYCDLSAETSDYEHQYYAKVKAIWGTKCSKWAESGRFYPFLETQI
GPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIYSNLKYNVSVLNTKSNRTWSQCVTN
HTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLKDQSSEFKAKIIFWYVLPISITV
FLFSVMGYSIYRYIHVGKEKHHPANLILYGNFDRFFVPAEKIVINFITLNISSDDSKISHQD
MSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFCDSEENTEGTSLTQQESLS
RTIPDPKTVIEYEDVRTTDICAGPEEQELSLQEEVSTQGTLLSQAALAVLGPQTLQYSYTP
QLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCEPSEGDGLGEE
GLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 240-260

N-glycosylation sites.amino acids 31-34, 72-75, 80-83, 171-174, 180-183, 189-192,
304-307, 523-526**Tyrosine kinase phosphorylation site.**

amino acids 385-392, 518-526

N-myristoylation sites.

amino acids 53-58, 106-111, 368-373, 492-497

Tissue factor

amino acids 1-278

189/550

FIGURE 189

ATGTGCTGCTGGCCGCTGCTCCTGCTGTGGGGGCTGCTCCCCGGGACGGCGGGGGGGGCTCG
GGCCGAACCTATCCGCACCGGACCCCTCCTGGACTCGGAGGGCAAGTACTGGCTGGGCTGGAGC
CAGCGGGGCAGCCAGATCGCCTTCCGCCTCCAGGTGCGCACTGCAGGCTACGTGGGCTTCCGGC
TTCTCGCCCCACCGGGGCCATGGCGTCCGCCGACATCGTCGTGGGCGGGGTGGCCCCACGGGCGG
CCCTACCTCCAGGATTATTTTACAAATGCAAATAGAGAGTTGAAAAAAGATGCTCAGCAAGAT
TACCATCTAGAATATGCCATGGAAAAATGACACACACAATAATTGAATTTACCAGAGAGCTG
CATAATGTGACATAAATGACAAGAGTATAACGGATAGCACTGTGAGAGTGATCTGGGCCTAC
CACCATGAAGATGCAGGAGAAGCTGGTCCCAAGTACCATGACTCCAATAGGGGCACCAAGAGT
TTGCGGTTATTGAATCCTGAGAAACTAGTGTGCTATCTACAGCCTTACCATACTTTGATCTG
GTAAATCAGGACGTCCCCATCCCAAACAAAGATACAACATATTGGTGCCAAATGTTTAAGATT
CCTGTGTTCCAAGAAAAGCATCATGTAATAAAGGTTGAGCCAGTGATACAGAGAGGCCATGAG
AGTCTGGTGCACCATCCTGCTCTATCAGTGCAGCAACAACCTTTAACGACAGCGTTCTGGAG
TCCGGCCACGAGTGCTATCACCCCAACATGCCCGATGCATTTCCTCACCTGTGAAACTGTGATT
TTTGCTGGGCTATTGGTGGAGAGGGCTTTTCTTATCCACCTCATGTTGGATTATCCCTTGGC
ATCCATTAGATCCGCATTATGTGCTCCTAGAAGTCCATTATGATAATCCCACTTATGAGGAA
GGCTTAATAGATAATTCTGGACTGAGGTATTTTACACAATGGATATAAGGAAATATGATGCT
GGGGTGATTGAGGCTGGCCTCTGGGTGAGCCTCTTCCATACCATCCCTCCAGGGATGCCTGAG
TTCCAGTCTGAGGGTCACTGCACCTTTGGAGTGCCTGGAAGAGGCTCTGGAAGCCGAAAAGCCA
AGTGGAATTCATGTGTTTGCTGTTCTTCTCCATGCTCACCTGGCTGGCAGAGGCATCAGGCTG
CGTCATTTTCGAAAAGGGAAGGAAATGAAATTACTTGCCTATGATGATGATTTTGACTTCAAT
TTCCAGGAGTTTCAGTATCTAAAGGAAGAACAACAATCTTACCAGGAGATAACCTAATTACT
GAGTGTGCTACAACACGAAAGATAGAGCTGAGATGACTTGGGGAGGACTAAGCACCAGGAGT
GAAATGTGTCTCTCATACCTTCTTTATTACCCAAGAATTAATCTTACTCGATGTGCAAGTATT
CCAGACATTATGGAACAACCTCAGTTCATTGGGGTTAAGGAGATCTACAGACCAGTCACGACC
TGGCCTTTCATTATCAAAAGTCCCAAGCAATATAAAAACCTTTCTTTCATGGATGCTATGAAT
AAGTTTAAATGGACTAAAAAGGAAGGTCTCTCCTTCAACAAGCTGGTCCTCAGCCTGCCAGTG
AATGTGAGATGTTCCAAGACAGACAATGCTGAGTGGTTCGATTCAAGGAATGACAGCATTACCT
CCAGATATAGAAAGACCCCTATAAAGCAGAACCTTTGGTGTGTGGCACGTCTTCTTCTCTTCC
CTGCACAGAGATTTCTCCATCAACTTGCTTGTTTGCTTCTGCTACTCAGCTGCACGCTGAGC
ACCAAGAGCTTGTGATCAAAAATTCTGTTGGACTTGACAATGTTTTCTATGATCTGAACCTGTC
ATTTGAAGTACAGGTTAAAGACTGTGTCCACTTTGGGCATGAAGAGTGTGGAGACTTTTCTTC
CCCATTTTCCCTCCCTCCTTTTTCTTTCCATGTTACATGAGAGACATCAATCAGGTTCTCTT
CTCTTTCTTAGAAATACCTGATGTTATATATACATGGTCAATAAAATAAACTGGCCTGACTT
AAGATAACCATTTTAAAAAATTGGGCTGTCATGTGGGAATAAAAGAATTCTTTCTTCTCTAAA
AAAAAAA

190/550

FIGURE 190

MCCWPLLLLWGLLPGTAAGGSGRTYPHRTLLDSEGKYWLGWSQRGSQIAFRLQVRTAGYVGFG
FSPTGAMASADIVVGGVAHGRPYLQDYFTNANRELKKDAQQDYHLEYAMENSTHTIIEFTREL
HTCDINDKSITDSTVRVIWAYHHEDAGEAGPKYHDSNRGTSKSLRLLNPEKTSVLSTALPYFDL
VNQDVPIPNKDTTYWCQMFKIPVFQEKHHVIKVEPVIQRGHESLVHHILLYQCSNNFNDSVLE
SGHECYHPNMPDAFLT CETVIFAWAIGGEGFSYPPHVGLSLGTPLDPHYVLLEVHYDNPTYEE
GLIDNSGLRLFYTMDIRKYDAGVIEAGLWVSLFHTIPPGMPEFQSEGHCTLECLEEAEAEKP
SGIHVFAVLLHAHLA GRGIRLRHFRKGKEMKLLAYDDDFDNFQEFQYLKEEQTILPGDNLIT
ECRYNTKDRAEMTWGGLSTRSEMCLSYLLYYPRINLTRCASIPDIMEQLQFIGVKEIYRPVTT
WPFIIKSPKQYKNLSFMDAMNKF KWKKEGLSFNKLVLSPVNVRC SKTDNAEWSIQGMTALP
PDIERPYKAEPLVCGTSSSSSLHRDFSINLLVCLLLL SCTLSTKSL

Signal peptide:

amino acids 1-18

Transmembrane domains:

amino acids 56-73, 378-393, 583-602

N-glycosylation sites.

amino acids 114-118, 247-251, 476-480, 517-521

N-myristoylation sites.amino acids 11-17, 15-21, 20-26, 45-51, 68-74, 79-85, 290-296,
316-322, 337-343, 342-348, 456-462, 534-540, 582-588**Copper type II, ascorbate-dependent monooxygenases proteins.**

amino acids 271-321, 422-474

191/550

FIGURE 191

GCTTCAGCTGAAGAAAGAGAGGAATGAAGCGCCTTCTGCTTCTGTTTTGTTCTTTATAACAT
TTTCTTCTGCATTTCCCTTAGTCCGGATGACGGAAAATGAAGAAAATATGCAACTGGCTCAGG
CATATCTCAACCAGTTCTACTCTCTTGAAATAGAAGGGAATCATCTTGTTCAAAGCAAGAATA
GGAGTCTCATAGATGACAAAATTCGGGAAATGCAAGCATTTTTTGGATTGACAGTGACTGGAA
AACTGGACTCAAACACCCTTGAGATCATGAAGACACCCAGGTGTGGGGTGCCTGATGTGGGCC
AGTATGGCTACACCTCCCTGGGTGGAGAAAATACAACCTCACCTACAGAATAATAAACTATA
CTCCGGATATGGCACGAGCTGCTGTGGATGAGGCTATCCAAGAAGGTTAGAAGTGTGGAGCA
AAGTCACTCCACTAAAATTCACCAAGATTTCAAAGGGGATTGCAGACATCATGATTGCCTTTA
GGACTCGAGTCCATGGTCCGTGTCCTCGCTATTTTGATGGTCCCTTGGGAGTGCTTGGCCATG
CCTTTCCTCCTGGTCCGGTCTGGGTGGTGACACTCATTTTGATGAGGATGAAACTGGACCA
AGGATGGAGCAGGATTCAACTTGTTTCTTGTGGCTGCTCATGAATTTGGTCATGCACTGGGGC
TCTCTCACTCCAATGATCAAACAGCCTTGATGTTCCCAAATTATGTCTCCCTGGATCCCAGAA
AATACCCACTTTCTCAGGATGATATCAATGGAATCCAGTCCATCTATGGAGGTCTGCCTAAGG
TACCTGCTAAGCCAAAGGAACCCACTATACCCCATGCCTGTGACCCTGACTTGACTTTTGACG
CTATCACAACCTTCCGCAGAGAAGTAATGTTCTTTAAAGGCAGGCACCTATGGAGGATCTATT
ATGATATCACGGATGTTGAGTTTGAATTAATTGCTTCATTCTGGCCATCTCTGCCAGCTGATC
TGCAAGCTGCATACGAGAACCCAGAGATAAGATTCTGGTTTTTAAAGATGAAAACCTTCTGGA
TGATCAGAGGATATGCTGTCTTGCCAGATTATCCCAAATCCATCCATACATTAGGTTTTCCAG
GACGTGTGAAGAAAATAGATGCAGCCGTCTGTGATAAGACCACAAGAAAACCTACTTCTTTG
TGGGCATTTGGTGCTGGAGGTTTGATGAAATGACCCAAACCATGGACAAAGGATTCCTCGCAGA
GAGTGGTAAAACACTTTCCTGGAATCAGTATCCGTGTTGATGCTGCTTTCCAGTACAAAGGAT
TCTTCTTTTTCAGCCGTGGATCAAAGCAATTTGAATACAACATTAAGACAAAGAATATTACCC
GAATCATGAGAACTAATACTTGGTTTCAATGCAAAGAACCAAGAACTCCTCATTTGGTTTTG
ATATCAACAAGGAAAAAGCACATTCAGGAGGCATAAAGATATTGTATCATAAGAGTTTAAGCT
TGTTTATTTTTGGTATTGTTCAATTTGCTGAAAAACACTTCTATTTATCAATTAAATTCATAGAC
CTAAAATAAACCTCAACAGGTCTTTTAATATAAATTCTGCTTCAAATAGATAAAACCATTC
TTTAACAAC

192/550

FIGURE 192

MKRLLLLFLFFITFSSAFPLVRMTENEENMQLAQAYLNQFYSLEIEGNHLVQSKNRS LI DDKI
REMQAFFGLT VTGKLD SNTLEIMKT PRCGVPDVGQYGYTLPGWRKY NLTYRI INYTPDMARAA
VDEAIQEGLEVWSKV TPLKFTKISKGIADIMIAFRTRVHGRCPRYFDGPLGVLGHA FPPGPGL
GGDTHFDEDENWTKDGAGFN LFLVAAHEFGHALGLSHSNDQTALMFPNYVSLDPRKYPLSQDD
INGIQSIYGG LPKVPAKPKEPTIPHACDPDLTFDAITTFRREVMFFKGRHLWRIYYDITDVEF
ELIASFWPSLPADLQAAYENPRDKILVFKDENFWMIRGYAVLPDYPKSIHTLGFPGRVKKIDA
AVCDKTTRKTYFFVGIWCWRFDEMTQMDKGFPQRVVKHFPGISIRVDAAFQYKGFFFFSRGS
KQFEYNIKTKNITRIMRTNTWFQCKEPKNSSFGFDINKEKAHSGGIKILYHKSLSLFIFGIVH
LLKNTSIYQ

Signal peptide:

amino acids 1-17

N-glycosylation sites.

amino acids 55-59, 110-114, 200-204, 452-456, 470-474, 508-512

N-myristoylation site.

amino acids 71-77, 205-211, 223-229

Hemopexin domain signature.

amino acids 171-202, 207-238, 318-334

Neutral zinc metallopeptidases, zinc-binding region signature.

amino acids 213-223

Matrixins cysteine switch.

amino acids 89-97, 207-238

193/550

FIGURE 193

CACAATCAGGTCCCATTTCTATAGATGGGGAAACTGAGGCTTGAGGTCACATAGGCGTCGTTCA
AGGCTGGTATACCTGCACCCTCTCCCATGTGAACAACATGGTTCTGGGTAATGGGGGCTGTCA
TCCAGTCTCCTCCCTGCCCCCTGCTGGTGCACCTTCCTGCCTCTGCTGGTGCACCTTTCTGCCCCCT
ACTGGTATATTTGCTGCCTCTGCTGGGGCGCTTCCTGCCTCGGCTGGTGTATCTCCTGCCCCCT
GCTGGTGCACCTTTCTGCCCCCGCTGATGCACTTCCTGCCTCTGCTGGTGCACCTTCCTGGCTCT
GCTGGCACACTTCCTGCCTCTGCTGGTGCACCTTCCTGGCTCTGCTGGCGCACTTTCTGCCCC
TGCTGGTGTATTTCTGCCCCCTGCTGGTGTACTTCCTTCCCCTGCTGGTGCACCTTCCTGCCTC
TGCTGGCGCACTTCTTGCCTCTCCAGGCCCTACCTAGCCTCTCCCTCTTATATATGGAAGTCT
TCCCAGTTCACTGACACTGGTAACAGGGACTCTGCTCTTGGTGTTGCTGTCTGCCCTGGGGAT
GGGCATCTGTGTCTTCCTTTACTACTGCTGGCTCAGGACCCAGAGCTTTGAAGCATGTCCAGA
TGCAGGTCCGGGCACCAGAGTCTAAGGAGCCCCCTACACCCACCAGATTTTCCAATAAAGAGA
TGTTACCA

194/550

FIGURE 194

MVLGNGGCHPVSSLPLLVHFLPLLVHFLPLLVYLLPLLGRFLPRLVYLLPLLVHFLPPLMHFL
PLLVHFLALLAHFLPLLVHFLALLAHFPAPAGVFPAPAGVLPSPAGALPASAGALLASPGPT

Signal peptide:

amino acids 1-39

N-myristoylation sites.

amino acids 4-10, 109-115, 116-122

Leucine zipper pattern.

amino acids 14-36, 16-38, 17-39, 21-43, 24-46, 28-50, 31-53,
35-57, 38-60, 42-64, 45-67, 49-71, 52-74, 56-78, 59-81, 63-85,
65-87, 66-88

195/550

FIGURE 195

[illegible]

196/550

FIGURE 196

MRRLTRRLVLPVFGVLWITVLLFFWVTKRKLEVPTGPEVQTPKPSDADWDDLWDQFDERRYLN
AKKWRVGDDPYKLYAFNQRESERISSNRAIPDTRHLRCTLVYCTDLPPTSIIITFHNEARST
LLRTIRSVLNRTPTHLIREIILVDDFSNDPDDCKQLIKLPKVKCLRNNERQGLVRSRIRGADI
AQGTTLTFLDSHCEVNRDWLQPLLHRVKEDYTRVVCVIDIINLDTFTYIESASELRGGFDWS
LHFQWEQLSPEQKARRLDPTPIRTPIIAGGLFVIDKAWFDYLGKYDMDMDIWGGENFEISFR
VWMCSSLEIVPCSRVGHVFRKKHPYVFPDGNANTYIKNTKRTAEVWMDEYKQYYYAARPFAL
ERPFGNVESRLDLRKNLRCQSFKWYLENIYPELSIPKESSIQKGNIRQRQKCLESQRQNNQET
PNLKLSPCAKVKGEDAKSQVWAFTYTQQILQEELCLSVITLFPGAPVVLVLCKNGDDRQQWTK
TGSHEHIAHLCLDTDMFGDGTENGKEIVVNPCESSLMSQHWDMVSS

Transmembrane domain:

amino acids 475-493

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 2-6

Tyrosine kinase phosphorylation sites.

amino acids 68-75, 401-409

N-myristoylation sites.

amino acids 178-184, 186-192, 192-198, 346-352, 383-389, 526-532

197/550

FIGURE 197

GCAGCTCACCCCTTCGACGCCGCGATGGGGGAAGACGACGCCGCGCTTCGGGCTGGCAGCAGGGGGCTCTCCGACC
CGTGGGCAGACTCAGTGGGAGTGGGACCCCGCACCACGGAGCGCCACATCGCCGTACACAAGCGGCTTGTGCTGG
CCTTCGCTGTGTCCCTCGTGGCATTGCTCGCGGTCACAATGCTCGCTGTGCTGCTCAGCCTGCGCTTCGACGAGT
GCGGGGCGAGTGCCACGCCAGGCGCGGACGGTGGCCCCCTCAGGCTTTCGGGAGCGCGGCGGCAACGGGAGCCTCC
CTGGATCGGCCCGCGCAACCACCACGACGGCGGGGACTCCTGGCAGCCCGAGGCGGGTGGGGTGGCCAGTCCGG
GGACCACGTGGGCCAGCCGCCGTGCGAGGAGGAGCGGGAGCCGTGGGAGCCGTGGACGCAGCTGCGCCTGTCCG
GCCACCTGAAGCCGCTGCACTACAATCTGATGCTCACCGCCCTTCATGGAGAACTTACCTTCTCCGGGGAGGTCA
ACGTGGAGATCGCGTGGCGAAGCCACCCGCTACGTAGTGTGCACGCTTCCCGAGTGGCGGTGGAGAAAGTGC
AGCTGGCCGAGGACCGGGCGTTCCGGGCTGTCCCTGTAGCCGTTTTTCTCTACCCGCAAACCCAGGTCTTAG
TGGTGGTGTGAATAGGACACTGGACGCGCAGAGGAATTACAATCTGAAGATTATCTACAACGCGCTCATCGAGA
ATGAGCTCCTGGGCTTCTTCCGACGCTCCTATGTGCTCCACGGGAGAGAAAGATTCTTGGTGTACTCAGTTTT
CGCTACACATGCCAGAAAGGCATTTCTTGTGTTGATGAGCCAATCTACAAGGCTACTTTCAAAATCAGCATCA
AGCATCAAGCAACCTATTTATCTTAATATGCCAGTGGAACTTCCGTGTTGAGGAAGATGGATGGGTTA
CGGATCACTTTTACAGACCCCTCATGTCCACATATTATTAGCCTGGGCAATTTGCAACTTCACATACAGAG
AAACTACCACCAAGAGTGGGGTGTAGTACGATTATATGCAAGACCTGATGCTATCAGAAGAGGATCCGGGGACT
ATGCTCTCCATATAACAAAGAGATTAATAGAATTTTATGAAGACTACTTTAAAGTGCCCTATTCTTGCCAAAAC
TAGATCTTTTAGCTGTGCCTAAGCATCCGTATGCTGCTATGGAGAAGTGGGGACTAAGTATTTTTGTGGAACAAA
GAATACTGCTGGATCCCAGTGTTCATCTATTTCTTATTTGCTGGATGTCACCATGGTCATTTGTCATGAGATAT
GTCACCAAGTGGTTGGTGACCTTGTGACGCTGTGTGGTGGGAAGACGTGTGGCTGAAGGAAGGGTTTGTCTACT
ACTTTGAATTTGTGGTACAGACTACCTCTATCCTGGCTGGAACATGGAAAAGCAGAGGTTTCTGACCGATGTTT
TGCATGAAGTGATGCTGCTGGACGGTTTGGCCAGTTCCTCATCCAGTATCAGAGGAAGTGTGTCAGGCAACAGATA
TTGACAGGGTGTGCTGAGTGGATCGCATATAAAAAGGGTGTGCTTTAATAAGAATGCTGGCTAATTTTATGGGCC
ATTGAGTTTCCAGAGGGTTTGAAGATTATTTAAACCATTTATAAGTATGGAATGTCAGCCAGAAATGATCTCT
GGAATACATTATCGGAGGCTTTAAAAAGAAATGGGAATATGTAATATACAAGAAGTAATGGATCAGTGGACAC
TCCAGATGGGTATCTCTGTTATCACCATCTTGGGAAACACAACAGCAGAAAATAGAATAATAATTACCCAACAGC
ATTTTATCTATGATATCAGTGTAAAACATAAGCACTTAAACTTCAGAAATACAGTTACCTGTGGCAGATTCCAT
TAACTATTGTGGTAGGAAATAGAAGCCATGTGTCTCAGAAGCAATTATTTGGGTGTCTAACAAATCAGAGCACC
ACAGAATAACTTATTTGGACAAAGGAAGCTGGCTGTGGGGAACATCAATCAAATGGCTATTTTAGAGTCAACT
ATGACCTAAGGAAGTGGAGATTATTAATTGATCAATTAATCCGGAATCATGAGGTTCTTTCTGTCTAGTAACCGAG
CGGGCTTGATCGATGATGCCTTACGCTAGCCAGGGCTGGCTATTTGCCTCAGAATATTCTCTGGAGATTATCA
GATACCTGTCTGAGGAGAAGGATTTTCTTCTTGGCATGCTGCCAGCCGAGCTCTTTATCTCTAGATAAATTAC
TGGACCGCATGGAAAACATAACATTTTCAATGAATATATTTTAAAGCAAGTTGCAACAACATATATCAAGCTTG
GGTGGCCGAAAAATAATTTTAAATGGATCTCTTGTTCAGCATCTTACCAACATGAAGAACTACGTAGAGAAGTTA
TAATGCTGGCTGCGAGTTTTGGCAACAAGCACTGTCCACCAACAGGCATCAACACTTATTTAGATTGGATTTCCA
GCAACAGGAACAGAATACCACTAAATGTTAGAGACATCGTATACTGTACAGGAGTGTCACTACTGGATGAGGATG
TCTGGGAATTCATATGGATGAAATTCATTCCACCACAGCAGTTTCTGAGAAGAAATATATTGGAAGCCTTAA
CTTGCAAGTGTGACAGGAATTTATTAACAGGCTTCTAAATCTGTCACTGAATTCTGAGGTGGTGTGGATCAAG
ATGCAATTGATGTATAATCCATGTAGCTCGAAATCCACATGGTTCGAGACCTTGCTGGAGTTTTCAGGGATA
AATGGAAGATATTAATACCAGGTATGGAGAAGCATTGTTTATGTATTCCAACTCATCAGTGGTGTACAGAAT
TTCTTAATACTGAAGGTGAAGTCAAGAGCTCAAGAACTTCATGAAAACTATGATGGGGTAGCTGCTGCTTCTT
TCTCAGGAGCTGTGGAAGCTGTGGAAGCAATGTGCGCTGGAAAATGCTTTACCAAGACGAGCTTTTCCAATGGT
TAGGAAAAGCTCTAAGACACTAAATATATGTATCTTATAAACAACAACTCAACTCAGAAGTTTATGAGAAGACAC
GCTTTTTTGTGGAATGAGGAAAATGTACTACCTAGAAAATGGCCAGATTTTCAGTGTAACTGTGGGAGGAATTT
TTTTTTTTAGTTTTATTTTTTGGTTTTTGGGGGATATTTTTTATTTGTTTCATTCTGTTCTGTTCTCTAC
TGGGTGTTCTCTCTAAAGAACTCTTGAAGTGAAGTACCATGATTGCTTCAGCTGTACATTCTTGTCTGTA
CAGGACCAATATGATAGTGATGATGTTGATGTTACAGTCAATTTGGAAAAACATATTGAGAATATCTGTGCAT
GGATATATTGCTGCTGTGTTCCAGCATGCTTATTTCAAACGTCCAGTGTGTGTGTAATATGTGTTACACC
TAGGATGGGCAATTATGCAAAAGCACAAGATTATATGACAATCAGTATTGCAATGAAAGAAAACTAAAAACA
GAAATGATATTCTCAATTTTGGCAATGTGAGAGGTAAAAATAGCCCTTGACATGATGAACATCACTTATTTACGC
ACTTGGATTGTCTGGCAATGATTACTGTGTTGCTAACTCATTTTCTTGGATTAAAGCTGTGTATACATTTTAAA
AGGCATATAGATAGTGATGATATGTATATGTACATAGGGAAGCCCATATGTATATAGTATGTTGTACACTGC
ACATGTACAAAGAATGTCTTCAGATCAAAGAAAATTTATCTCTTTTATAAACTTAAGGACAGTTGCAAAAGGCT
TCAAGGAATTTATCTCAACATATTCTTTCTATGCTCACTAAATTTCTCACTGTATGAATTTTTCTATCTAC
TTCTTGAACAGTCTATTCTGTACATGAAGATGAATACAAACAAAATTTTTGTATAAACTCCCAAAAAA
AAAAA

198/550

FIGURE 198

MGEDDAALRAGSRGLSDPWADSVGVRPRRTTERHIAVHKRLVLAFAVSLVALLAVTMLAVLLSL
RFDECGASATPGADGGPSGFPERGGNGSLPGSARRNHHAGGDSWQPEAGGVASPGTTSAQPPS
EEEREPWEPWTQLRLSGHLKPLHYNLMLTAFMENFTFSGEVNVEIACRNATRYVVLHASRVAV
EKVQLAEDRAFGAVPVAGFFLYPQTQVLVVVLNRTLDAQRNYNLKKIYNALIENELLGFFRSS
YVLHGERRFLGVTQFSPTHARKAFPCFDEPIYKATFKISIKHQATYLSLSNMPVETSVFEEDG
WVTDHFSQTPLMSTYYLAWAICNFTYRETTTKSGVVRLYARPDARRGSGDYALHITKRLIE
FYEDYFKVPYSLPKDLLAVPKHPYAAMENWGLSIFVEQRILLDPSVSSISYLLDVTMVIVHE
ICHQWFGDLVTPVWWEDVWLKEGFAHYFEFVGTDYLYPGWNMEKQRFLTDLVHEVMLLDGLAS
SHPVSQEVQLQATDIDRVFDWIAYKKGAALIRMLANFMGHSVFQRLQDYLTIHKYGNAARNDL
WNTLSEALKRNGKYVNIQEVMDQWTLQMGYPVITILGNTTAENRIIITQQHFIYDISAKTKAL
KLQNNSYLWQIPLTIVVGNRSHVSSEAIWVSNKSEHHRITYLDKGSWLLGNINQTYFRVNY
DLRNWRLIDQLIRNHEVLSVSNRAGLIDDAFSLARAGYLPQNIPLEIIRYLSEEKDFLPWHA
ASRALYPLDKLLDRMENYNIFNEYILKQVATTYIKLGWPKNNFNGLSVQASYQHEELRREVIM
LACSEFGNKHCHQOASTLISDWISSNRNRIPLNVRDIVYCTGVSLLEDVWEFIWMKFHSTTAV
SEKKILLEALTCSDDRNLLNRLNLSLNSEVVDQDAIDVIIHVARNPHGRDLAWKFFRDKWK
ILNTRYGEALFMYSKLISGVTEFLNTEGELKELKNFMKNYDGVAAASFSRAVETVEANVRWKM
LYQDELFWLQWLGKALRH

Transmembrane domain:

amino acids 44-63

N-glycosylation sites.amino acids 89-93, 160-164, 175-179, 222-226, 338-342, 605-609,
634-638, 649-653, 663-667, 684-688, 800-804, 906-910**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 362-366

Tyrosine kinase phosphorylation site.

amino acids 520-528

N-myristoylation sites.amino acids 78-84, 87-93, 90-96, 118-124, 501-507, 604-610,
825-831, 987-993**Neutral zinc metallopeptidases, zinc-binding region signature.**

amino acids 437-447

199/550

FIGURE 199

GGCCCCGGCGCAGCTCGGCCAGAGCGACCGGGGGCTGAGCGCGCGTCCGCCAGGGGGCTCCGGAAGCTGCCCC
GGCCCGCGGCTCCTCCCTCGTCCCGCTTCCCTTTCTCGCTCACCGCCGCCCTCCTTCCCCAGCTCCCTCGCC
GTCCGCCCGCCCCACAGCCAGCGGCTCCCGCCCCCTGCAGCCACGATGCCCCGGCCCCGGCCGCCCGCGGG
ACTCCGCGGGATCTCGCTGTTCTCGCTCTGCTCCTGGGGAGCCCCGGCGGCAGCGCTGGAGCGAGATGCTCTTCC
CGAGGGAGATGCTAGCCCTTTGGGTCTTACCTCCTGCCCTCAGGAGCCCCGGAGAGAGCGAGTCTGGCAAAGA
GCACCCTGAAGAGAGAGTGGTAACAGCGCCCCCAGTTTCTCACAGTCGGCGGAAGTGTGGGCGAGCTGGTGCT
GGATGGGACCGCACCTCTGCACATCACGACATCCAGCCCTGTACCCTGCTTCCAGAGGAGGCCCGCCCCAA
GCACGCCCTTGCCCCCAAGAAGAACTGCCCTTCGCTCAAGCAGGTGAACTCTGCCAGGAAGCAGCTGAGGCCCAA
GGCCACCTCCGCAGCCACTGTCCAAAGGGCAGGGTCCCAGCCAGCGTCCCAGGGCCTAGATCTCCTCTCCTCCTC
CACGGAGAAGCCTGGCCACCGGGGACCGGACCCCATCGTGCCCTCCGAGGAGGCATCAGAAGTGCCCTTTG
GCTGGATCGAAAGGAGAGTGGGTCCCTACAACACCCGCACCCCTGCAAACTCTCCCCCTTCACTTCGCAGCCCTA
TGTGGCCACACACTCCCCCAGAGGCCAGAACCCGGGGAGCCTGGGCTGACATGGCCAGGAGGCCCGCCCAA
GGACACCGCCCATGGCCCTGATGGACAAAGGTGAGAATGAGCTGACTGGGTGAGCCTCAGAGGAGAGCCAGGA
GACCACTACCTCCACCATTATCACCACCAGGTCATCACCACCGAGCAAGCACCAGCTCTCTGCAGTGTGAGCTT
CTCCAATCTGAGGGGTACATTGACTCCAGCGACTACCCACTGCTGCCCTCAACAACTTTCTGGAGTGACATA
CAACGTGACACTCTACACTGGCTATGGGGTGGAGCTCCAGGTGAAGAGTGTGAACCTGTCCGATGGGGAAGTGT
CTCCATCCGCGGGGTGGACGGCCCTACCTGACCGTCTTGCCCAACAGACACTCCTGGTGGAGGGGCAGGTAAT
CCGAAGCCCCACCAACACCATCTCCGTCTACTTCCGGACCTTCCAGGACGACGGCCTTGGGACCTTCCAGCTTCA
CTACCAGGCCTTCATGCTGAGCTGCAACTTTCCCCCGCGGCCTGACTCTGGGGATGTACGGTGTGGACCTGCA
CTCAGGTGGGTGGCCACTTTCAGTCCACCTGGGCTATGAGCTCCAGGGCGCTAAGATGCTGACATGCATCAA
TGCTCCAGCCGCACTGGAGCAGCCAGGAGCCCATCTGCTCAGCTCCTTGTGGAGGGGCAGTGCACAATGCCAC
CATCGGCCGCGTCTCTCCCAAGTTACCCTGAAAACACAAATGGGAGCCAATTCTGCATCTGGACGATTGAAGC
TCCAGAGGGCCAGAGCTGCACCTGCACCTTGGAGAGGCTGTTGCTGCATGACAAGGACAGGATGACGGTTACAG
CGGGCAGACCAACAAGTCAGCTCTTCTCTACGACTCCCTTCAAACCGAGAGTGTCCCTTTTGGGGCCTGCTGAG
CGAAGGCAACACCATCCGCATCGAGTTACGTCGACCGAGGCCCGGGCGGCCTCCACCTTCAACATCCGATTTGA
AGCGTTTGAAGAGGCCACTGCTATGAGCCCTACATCCAGAATGGGAAGTTCCTACATCCGACCCGACCTATAA
CATGGGACTATAGTGGAGTTACCTGCGACCCCGCCACTCCCTGGAGCAGGGCCCGGCCATCATCGAATGCAT
CAATGTGCGGGACCCATACTGGAATGACACAGAGCCCTGTGTCAGAGCCATGTGTGGTGGGGAGCTCTCTGCTGT
GGCTGGGGTGGTATTGTCCCCAACTGGCCCGAGCCCTACGTGGAAGGTGAAGATTGTATCTGGAAGATCCACGT
GGGAGAAGAGAAACGGATCTTCTTAGATATCCAGTTTCTGAATCTGAGCAACAGTGACATCTTGACCATCTACGA
TGGCGACGAGGTCATGCCCCACATCTTGGGGCAGTACCTTGGGAACAGTGGCCCCCAGAACTGTACTCTCCAC
GCCAGACTTAACCATCCAGTTCCATTCCGACCTGCTGGCCTCATCTTTGGAAAGGGCCAGGGATTTATCATGAA
CTACATAGAGGTATCAAGGAATGACTCCTGCTCGGATTTACCCGAGATCCAGAATGGCTGGAAACCACTTCTCA
CACGGAGTTGGTGGGGGAGCCAGAATCACCTACCAGTGTGACCCCGGCTATGACATCGTGGGGAGTGACACCT
CACCTGCCAGTGGGACCTCAGCTGGAGCAGCGACCCCCATTTGTGAGAAAATTATGTACTGCACCGACCCCGG
AGAGGTGGATCACTCGACCCGCTTAATTTCCGATCCTGTGCTGCTGGTGGGGACCACCATCCAATACACCTGCAA
CCCCGTTTGTGCTTGAAGGGAGTTCTTCTGACCTGCTACAGCCGTGAACAGGGACTCCCATCTGGACGTC
TCGCCTGCCCACTGCGTTTCGGAGGAGTCCCTGGCATGTGACAACCCAGGGCTGCCTGAAAATGGATACCAAT
CCTGTACAAGCGACTCTACCTGCCAGGAGTCCCTCACCTTCATGTGCTACGAAGGCTTTGAGCTCATGGGTGA
AGTGACCATCCGCTGCATCCTGGGACAGCATCCCACTGGAACGGGCCCCCTGCCGTGTGTAAAGTTAATCAAGA
CAGTTTTGAACATGCTTTAGAAGCAGAAGCGGCAGCAGAGACGTCGCTGGAAGGGGGGAACATGGCCCTGGCTAT
CTTCATCCCGGTCTCATCTCTCTTACTGCTGGGAGGAGCTACATTTACATCACAAGATGTCGCTACTATTTC
CAACCTCCGCCTGCCTCTGATGTACTCCACCCCTACAGCCAGATCACCGTGGAACCGAGTTTGACAACCCCAT
TTACGAGACAGGGGAACAGAGAGTATGAGGTTTCTATCTAAAGAGAGCTACACTTGAGAAGGGGACTTGTGAA
CTCAACCACAATCTCCTCGAGACATTCATCCAGAGACCATGTGGCACTTGATTGAAACCCAGAAATGTCGACTGT
CTTTTGTGTAGACTCTTTATCAAAGGTTTACTGTTTCTTCCTGTATTTATTATATTTAAAGTGAAAAAAA
AAAAAAA

200/550

FIGURE 200

MPAARPPAAGLRGISLFLALLLGSPAAALERDALPEGDASPLGPYLLPSGAPERGSPGKEHPE
ERVVTAPPSSSSQSAEVLGELVLDGTAPSAHHDIPALSPLLPEEARPKHALPPKKKLPSLKQVN
SARKQLRPKATSAATVQRAGSQPASQGLDLLSSSTEKPGPPGDPDPIVASEEASEVPLWLDRK
ESAVPTTPAPLQISPFQTSQPYVAHTLPQRPEPGEPPDMAQEAPQEDTSPMALMDKGENELTG
SASEESQETTTSTIITTTVITTEQAPALCSVSFSNPEGYIDSSDYPLLPLNNFLECTYNVTYV
TGYGVELQVKSVNLSDGELLSIRGVDGPTLTVLANQTLLEGGQVIRSPNTNISVYFRTFQDDG
LGTFLHYQAFMLSCNFPRRPDSGDVTVMDLHSGGVAHFHCHLGYELQGAKMLTCINASKPHW
SSQEPICSAPCGGAVHNATIGRVLSPSPYENTNGSQFCIWTIEAPEGQKLHLHFERLLLHDKD
RMTVHSGQTNKSALLYDSLQTESVPFEGLLSEGNTIRIEFTSDQARAASTFNIRFEAFEKGHC
YEPYIQNGNFTTSDPTYNIGTIVEFTCDPGHSLEQGPPIECINVRDPYWNDTEPLCRAMCGG
ELSAVAGVVLSPNWPEPYVEGEDCIWKIHVGEEKRIFLDIQFLNLSNSDILTIYDGDVEMPHI
LGQYLGNSGPQKLYSSTPDLTIQFHSDPAGLIFGKGQGFIMNYIEVSRNDSCSDLPEIQNGWK
TTSHTELVRGARITYQCDPGYDIVGSDTLTCQWDLWSDDPPFCEKIMYCTDPGEVDHSTRLI
SDPVLLVGTTIQTTCNPGFVLEGSSLLTCYSRETGTPIWTSRLPHCVSEESLACDNPGLPENG
YQILYKRLYLPGESLTFMCYEGFELMGEVTIRCILGQPSHWNGPLPVCKVNQDSFEHALEAEA
AAETSLEGGMALAIPIVLIISLLLGGAYIYITRCRYYSNLRPLMYSHPYQSITVETEFDN-
PIYETGETREYEYSI

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 893-915

N-glycosylation sites.amino acids 311-315, 328-332, 350-354, 435-439, 458-462, 474-478,
514-518, 576-580, 618-622, 674-678, 742-746**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 188-192

N-myristoylation sites.amino acids 23-29, 87-93, 146-152, 454-460, 475-481, 575-581,
629-635, 695-701, 723-729, 766-772, 877-883, 953-959**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 383-394

201/550

FIGURE 201

GATGGCTACGGCAGGGGGTGGCTCTGGGGCTGACCCGGGAAGTCGGGGTCTCCTTCGCCTTCT
GTCTTTCTGCGTCCTACTAGCAGGTTTGTGCAGGGGAACTCAGTGGAGAGGAAGATATATAT
CCCCTTAAATAAAACAGCTCCCTGTGTTTCGCCTGCTCAACGCCACTCATCAGATTGGCTGCCA
GTCTTCAATTAGTGGAGACACAGGGGTATCCACGTAGTAGAGAAAGAGGAGGACCTACAGTG
GGTATTGACTGATGGCCCCAACCCCTTACATGGTTCTGCTGGAGAGCAAGCATTTTACCAG
GGATTTAATGGAGAAGCTGAAAGGGAGAACCAGCCGAATTGCTGGTCTTGCAGTGTCCTTGAC
CAAGCCCAGTCCTGCCTCAGGCTTCTCTCCTAGTGTACAGTGCCCAAATGATGGGTTTGGTGT
TACTCCAATTCCATATGGGCCAGAGTTTGTCTACTGCAGAGAAATACAGTGGAAATTCGCTGGG
CAATGGTTTGGCTTATGAAGACTTTAGTTTCCCCATCTTTCTTCTTGAAGATGAAAATGAAAC
CAAAGTCATCAAGCAGTGCTATCAAGATCACAACCTGAGTCAGAATGGCTCAGACCAACCTT
CCCCTATGTGCCATGCAGCTCTTTTACACATGCATGCTGTCATCAGCACTGCCACCTGCAT
GCGGCGCAGCTCCATCCAAAGCACCTTCAGCATCAACCCAGAAATCGTCTGTGACCCCTGTCT
TGATTACAATGTGTGGAGCATGCTAAAGCCTATAAATACAACCTGGGACATTAAGCCTGACGA
CAGGGTTGTGGTTGCTGCCACCCGGCTGGATAGTCGTTCTTTTCTGGAATGTGGCCCCAGG
GGCTGAAAGCGCAGTGGCTTCTTTGTCAACCCAGCTGGCTGCTGCTGAAGCTTTGCAAAAGGC
ACCTGATGTGACCACCCTGCCCGCAATGTCATGTTGTCTTCTTTCAAGGGGAACTTTTGA
CTACATTGGCAGCTCGAGGATGGTCTACGATATGGAGAAGGGCAAGTTTCCCGTGCAGTTAGA
GAATGTTGACTCATTTGTGGAGCTGGGACAGGTGGCCTTAAGAACTTCATTAGAGCTTTGGAT
GCACACAGATCCTGTTTCTCAGAAAAATGAGTCTGTACGGAACCAGGTGGAGGATCTCCTGGC
CACATTGGAGAAGAGTGGTGTGCTGGTGTCCCTGCTGTCATCCTCAGGAGGCCAAATCAGTCCCA
GCCTCTCCCACCATCTTCCCTGCAGCGATTTCTTCGAGCTCGAAACATCTCTGGCGTTGTTCT
GGCTGACCACTCTGGTGCCTTCCATAACAAATATTACCAGAGTATTTACGACACTGCTGAGAA
CATTAATGTGAGCTATCCCGAATGGCTGAGCCCTGAAGAGGACCTGAACCTTTGTAACAGACAC
TGCCAAGGCCCTGGCAGATGTGGCCACGGTGTGGGACGTGCTCTGTATGAGCTTGCAGGAGG
AACCAACTTCAGCGACACAGTTCAGGCTGATCCCCAAACGGTTACCCGCCTGCTCTATGGGTT
CCTGATTAAAGCCAACAACCTCATGGTTCCAGTCTATCCTCAGGCAGGACCTAAGGTCCTACTT
GGGTGACGGGCCTCTTCAACATTACATCGCTGTCTCCAGCCCCACCAACACCCTTATGTTGT
ACAGTATGCCTTGGCAAATTTGACTGGCACAGTGGTCAACCTCACCCGAGAGCAGTGCCAGGA
TCCAAGTAAAGTCCCAAGTGAACAAGGATCTGTATGAGTACTCATGGGTCCAGGGCCCTTT
GCATTCTAATGAGACGGACCGACTCCCCCGGTGTGTGCGTTCTACTGCACGATTAGCCAGGGC
CTTGTCTCCTGCCTTTGAACTGAGTCAGTGGAGCTCTACTGAATACTCTACATGGACTGAGAG
CCGCTGGAAGATATCCGTGCCCCGATATTTCTCATCGCCAGCAAAGAGCTTGAGTTGATCAC
CCTGACAGTGGGCTTCGGCATCCTCATCTTCTCCCTCATCGTCACCTACTGCATCAATGCCAA
AGCTGATGTCTTTTTCATTGCTCCCCGGGAGCCAGGAGCTGTGTCATAC**TGAGG**AGGACCCCA
GCTTTTCTTGCCAGNTCAGCAGTTCACCTTCCTAGAGCATCTGTCCCCTGGGACACAACCCT
AATTTGTCACTGGAACCTCCCTGGGCCTGTCTCAGATTGGGATTAAACATAAAAGAGTGGAAC
ATCCAAAAGAGACAGGGAGAAATAAATAAATTGCCTCCCTTCCCTCCGCTCCCCTTTCCCATCA
CCCCCTCCCCATTTCCCTCTTCTCTCTACTCATGCCAGATTTTGGGATTACAAATAGAAGCT
TCTTGCTCCTGTTTAACTCCCTAGTTACCCACCCTAATTTGCCCTTCAGGACCCTTCTACTTT
TTCCTTCCCTGCCCTGTACCTCTCTCTGCTCCTCACCCCCACCCCTGTACCCAGCCACCTTCCCT
GACTGGGAAGGACATAAAAGGTTTAAATGTCAGGGTCAAACCTACATTGAGCCCCTGAGGACAGG
GGCATCTCTGGGCTGAGCCTACTGTCTCCTTCCCACTGTCCTTTCTCCAGGGCCCTCAGATGGC
ACATTAGGGTGGGCGTGCTGCGGGTGGGTATCCACCTCCAGCCCACAGTGCTCAGTTGTACT
TTTTATTAAGCTGTAATATCTATTTTGTGTTTTTGTCTTTTTCCTTTATTTCTTTTGTAAATAT
ATATATAATGAGTTTCATTAAAATAGATTATCCC

202/550

FIGURE 202

MATAGGGSGADPGSRGLLRLLSFCVLLAGLCRGNSVERKIYIPLNKTAPCVRLLNATHQIGCQ
SSISGDTGVIHVVEKEEDLQWVLTGPNPPYMVLLSKHFTRDLMEKLGRTSRIAGLAVSLT
KPSPASGFSPSVQCPNDGFGVYSNSYGPEFAHCREIQWNSLGNGLAYEDFSFPFLLEDENET
KVIKQCYQDHNLSQNGSAPTFLCAMQLFSMHMAVISTATCMRRSSIQSTFSINPEIVCDPLS
DYNVWSMLKPINTTGTCLKPDDRVVVAATRLDSRSFFWNVAPGAESAVASFVTQLAAAEALQKA
PDVTTLPRNVMFVFFQGETFDYIGSSRMVYDMEKGKFPVQLENVDSFVELGQVALRTSLELWM
HTDPVSQKNESVRNQVEDLLATLEKSGAGVPAVILRRPNQSQPLPSSLQRFLRARNISGVVL
ADHSGAFHNKYYQSIYDTAENINVSYPEWLSPEEDLNFVTD TAKALADVATVLGRALYELAGG
TNFSDTVQADPQTVTRLLYGFLIKANNSWFQSILRQDLRSYLG DGPLQHYIAVSSPTNTTYVV
QYALANLTGTVVNLTREQCQDPSKVPSENKDLYEYSWVQG PLHSNETDRLPRCVRSTARLARA
LSPAFELSQWSSTEYSTWTESRWKDIRARIFLIASKELELITLTVGFGILIFSLIVTYCINAK
ADVLFIAPREPGAVSY

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 671-692

N-glycosylation sites.

amino acids 45-49, 55-59, 187-191, 200-204, 204-208, 264-268,
387-391, 417-421, 435-439, 464-468, 506-510, 530-534, 562-566,
573-577, 580-584, 612-616

Glycosaminoglycan attachment site.

amino acids 404-408

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 232-236

N-myristoylation site.

amino acids 5-11, 6-12, 9-15, 29-35, 61-67, 120-126, 146-152,
168-174, 205-211, 294-300, 438-444, 446-452, 504-510, 576-582

203/550

FIGURE 203

GCTAGACCGAGCCCTGGGAGGCTACGGGCTCCCCCGGAAACCCTGCCAGGGGAGCCGGGTTTT
GAGCTCAGGCGCCTCTAGCGGCGGCCCCAGAAATCTGACTCGCGAGGCCAGAGTTGCAGGGA
CTGAATAGCAAACCTGAGGCTGAGTAGGGAACAGACC**ATG**AGGTGAGTGCAGATCTTCCTCTCC
CAATGCCGTTTGTCTCTTCTACTAGTTCGGACAATGCTCCTTAAGTCTCTTGGCGAAGATGTA
ATTTTTCACCCTGAAGGGGAGTTTGACTCGTATGAAGTCACCATTCCCTGAGAAGCTGAGCTTC
CGGGGAGAGGTGCAGGGTGTGGTCAGTCCCGTGTCTTACCTACTGCAGTTAAAAGGCAAGAAG
CACGTCCCTCCATTTGTGGCCCAAGAGACTTCTGTTGCCCCGACATCTGCGCGTTTTCTCCTTC
ACAGAACATGGGGAACCTGCTGGAGGATCATCCTTACATACCAAAGGACTGCAACTACATGGGC
TCCGTGAAAGAGTCTCTGGACTCTAAAGCTACTATAAGCACATGCATGGGGGGTCTCCGAGGT
GTATTTAACAATTGATGCCAAACATTACCAAATTGAGCCCCCTCAAGGCCTCTCCCAGTTTTTGAA
CATGTCGTCTATCTCCTGAAGAAAGAGCAGTTTGGAATCAGGTTTGTGGCTTAAGTGATGAT
GAAATAGAATGGCAGATGGCCCCCTTATGAGAATAAGGCGAGGCTAAGGGACTTTCTGGATCC
TATAAACACCCCAAAGTACTTGAATTGATCCTACTCTTTGATCAAAGTAGGTATAGGTTTTGTG
AACAAACAATCTTTCTCAAGTCATACATGATGCCATTCTTTTGAAGTGGGATTATGGACACCTAC
TTTCAAGATGTTTCGTATGAGGATACACTTAAAGGCTCTTGAAGTATGGACAGATTTTAAACAA
ATACGCGTTGGATATCCAGAGTTAGCTGAAGTTTTAGGCAGATTTGTAATATATAAAAAAAGT
GTATTAAATGCTCGCCTGTCATCAGATTGGGCACATTTATATCTTCAAAGAAAATATAATGAT
GCTCTTGATGGTCGTTTGGAAAAGTGTGTTCTCTAGAATATGCTGGATCAGTGAGTACTTTA
CTAGATACAAATATCCTTGCCCCCTGCTACCTGGTCTGCTCATGAGCTGGGTCATGCTGTAGGA
ATGTCACATGATGAACAATACTGCCAATGTAGGGGTAGGCTTAATTGCATCATGGGCTCAGGA
CGCACTGGGTTTAGCAATTGCAGTTATATCTCTTTTTTTAAACATATCTCTTCGGGAGCAACA
TGTCTAAATAATATCCCAGGACTAGGTTATGTGCTTAAGAGATGTGGAAACAAAATTGTGGAG
GACAATGAGGAATGTGACTGTGGTTCCACAGAGGAGTGTGAGAAAGATCGGTGTTGCCAATCA
AATTGTAAGTTGCAACCAGGTGCCAATGTAGCATTGGACTTTGCTGTCATGATTGTGCGGTTT
CGTCCATCTGGATACGTGTGTAGGCAGGAAGGAAATGAATGTGACCTTGCAGAGTACTGCGAC
GGGAATTCAAGTTCTTGCCCCAATGACGTTTATAAGCAGGATGGAACCCCTTGCAAGTATGAA
GGCCGTTGTTTCAGGAAGGGGTGCAGATCCAGATATATGCAGTGCCAAAGCATTTTTGGACCT
GATGCCATGGAGGCTCCTAGTGAGTGCTATGATGCAGTTAACTTAATAGGTGATCAATTTGGT
AACTGTGAGATTACAGGAATTCGAAATTTTAAAAAGTGTGAAAGTGCAAAATCAATATGTGGC
AGGCTACAGTGTATAAATGTTGAAACCATCCCTGATTGGCCAGAGCATACGACTATAATTTCT
ACTCATTTACAGGCAGAAAATCTCATGTGCTGGGGCACAGGCTATCATCTATCCATGAAACCC
ATGGGAATACCTGACCTAGGTATGATAAATGATGGCACCTCCTGTGGAGAAGGCCGGGTATGT
TTTTAAAAAAATTGCGTCAATAGCTCAGTCCTGCAGTTTGACTGTTTGCTGAGAAATGCAAT
ACCCGGGGTGTGTTGCAACAACAGAAAAAAGTCCACTGCATGTATGGGTGGGCACCTCCATTTC
TGTGAGGAAGTGGGGTATGGAGGAAGCATTGACAGTGGGCCTCCAGGACTGCTCAGAGGGGCG
ATTCCCTCGTCAATTTGGGTTGTGTCCATCATAATGTTTCGCCTTATTTTATTAATCCTTTCA
GTGGTTTTTGTGTTTTTCCGGCAAGTGATAGGAAACCACTTAAAACCCAAACAGGAAAAAATG
CCACTATCCAAAGCAAAAACCTGAACAGGAAGAATCTAAAACAAAAACTGTACAGGAAGAATCT
AAAACAAAAAAGTGGACAGGAAGAATCTGAAGCAAAAACCTGGACAGGAAGAATCTAAAGCAAAA
ACTGGACAGGAAGAATCTAAAGCAAACTTGAAGTAACGACCCAAAGCAAAGAGTGTCAAG
AAACAAAAAAG**TAA**ACCGGGCAATCCATACTCATTTCAGTAACACAGGCTCATTTATTTAACCA
GCTAATCATTTATCCAAAGGCTTTCCATTCTTCTCCCAATATTTTTTTTACTTTAATTTTCCC
ACAAGTTTTGATCAGCAAATAAACAGCATTCTTGTGTTTGGAAACAAAA

204/550

FIGURE 204

MRSVQIFLSQCRLLLLLLVPTMLLKSLGEDVIFHPEGEFDSYEVTIPEKLSFRGEVQGVVSPVS
YLLQLKGKKHVLHLWPKRLLLPRHLRVFSFTEHGELLEDPYIPKDCNYMGSVKESLDSKATI
STCMGGLRGVFNIDAKHYQIEPLKASPSFEHVYLLKKEQFGNQVCGLSDDEIEWQMAPYENK
ARLRDFPGSYKHPKYLELILLFDQSRVFNNNLSQVIHDAILLTGIMDTYFQDVRMRIHLKA
LEVWTFDNKIRVGYPELAEVLGRFVIYKKSVLNARLSSDWAHLYLQRYNDALAWSFGKVCSL
EYAGSVSTLLDTNILAPATWSAHELGHAVGMSHDEQYCQCRGRLNCIMGSGRTGFSNCSYISF
FKHISSGATCLNNIPGLGYVLKRCGNKIVEDNEECDGSTEECQKDRCCQSNCKLQPGANCSI
GLCCHDCRFRPSGYVCRQEGNECDLAEYCDGNSSSCPNDVYKQDGTTPCKYEGRCFRKGCRSRY
MQCQSIFGPDAMEAPSECYDAVNLIQDQFGNCEITGIRNFKKCESANSICGRLQCINVETIPD
LPEHTTIISTHLQAENLMCWGTGYHLSMKPMGIPDLGMINDGTSCGEGRVCFKKNCVNSSVLQ
FDCLPEKCNTRGVCNNRKNCHCMYGWAPPFCEEVGYGGSIDSGPPGLLRGAIPSSIWVSIIM
FRLILLILSVVFVFFRQVIGNHLKPKQEKMPLSKAKTEQEESEKTKTVQEESEKTKTGQEESEAK
TGQEESEKAKTGQEESEKANIIESKRPKAKSVKKQKK

Signal peptide:

amino acids 1-27

Transmembrane domain:

amino acids 684-705

N-glycosylation sites.

amino acids 222-226, 372-376, 438-442, 473-477, 625-629

N-myristoylation sites.amino acids 131-137, 168-174, 235-241, 319-325, 364-370, 436-442,
472-478, 609-615, 642-648, 668-674, 676-680, 680-686, 749-755,
758-764, 767-773**Amidation site.**

amino acids 69-73

Disintegrins proteins

amino acids 429-479

EGF-like domain proteins

amino acids 650-662

Neutral zinc metallopeptidases, zinc-binding region proteins

amino acids 335-345

205/550

FIGURE 205

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGGGAAGGTTGAATGGGGTAGAAGGCCTG
TTGTGGAGGGAAACCACCCATCCTCCTGCCTCCCACCACCACCATCATCCTGGCTGGACGGAG
AGGGTGACGGGGGCTGGGAAGGGGCAGCTCATGTTTCAAGGTTCCAGGAGGGGCTACCTGTTGA
CTGTCTTTGCAGGAAGAAGAAAACACCTGAGTGACCAGATGTCCAGCTCCAGGTGCCTTGCC
AGATGGCCAGAACCACACCTCTTGAAGAGTGACAGTGCTGTGGAGCATGGTTTCTGCACACCT
GGAATGACTGGAACCCCAAAGACTCAAGAAGGAGCTAAAGATCTTGAAGTAGACATGAATAAA
ACAGAAGGCTGTGGACCACCTGTCGAGATGGAGAAGTCCTTCTGAGGCTATCCAAACACGGAC
CAGGCCATGAGACCCCCGATGACCATCCCTGAATTTTTTTCGAGAGTCAGTCAACCGATTTGGAA
CTTATCCAGCCCTCCCATCCAAGAATGGCAAAAAGTGGGAAATCTGAATTTCAACCGACTACT
ATGAGGCTTGTGCGAAGGCTGCAAAATCCTTGATCAAGCTGGGTTTGGAGCGTTTCCACGGAG
TTGGTATCCTGGGGTTTAACTCTGCAGAGTGGTTTATCACTGCTGTTGGTGCCATCCTAGCCG
GGGGTCTTTGTGTTGGTATTTATGCCACCAACTCTGCCGAGGCTTGTCAATATGTCATCACTC
ATGCCAAAGTGAACATCTTGCTGGTTGAGAATGATCAACAGTTACAGAAAATCCTTTCGATTC
CACAGAGCAGCCTAGAGCCCCCTAAAAGCGATCATCCAGTACAGACTGCCAATGAAGAAGAACA
ACAACTTGACTCTTGGGATGATTTTATGGAACCTGGCAGAAGTATCCCTGACACCCAACTGG
AGCAGGTCATCGAGAGCCAGAAGGCGAATCAATGCGCAGTGCTCATCTACACTTCAGGGACCA
CAGGCATACCCAAGGGAGTGATGCTCAGTCATGACAACATCACGTGGATTGCAGGAGCAGTGA
CAAAGGACTTTAACTGACAGACAAGCATGAGACGGTGGTTAGCTACCTCCCACTCAGCCATA
TTGCAGCACAGATGATGGACATCTGGGTACCCATAAAGATTGGGGCGCTCACATACTTTGCTC
AAGCAGATGCTCTCAAGGGCACCTTGGTAAGTACTCTAAAGGAGGTAAAACCTACTGTCTTCA
TTGGAGTGCCTCAAATTTGGGAGAAGATACATGAGATGGTGAAGAAAAATAGTGCCAAGTCCA
TGGGCTTGAAGAAGAAGGCATTCTGTGGGCAAGAAACATTGGCTTCAAGGTCAACTCAAAAA
AGATGTTGGGGAAATATAATACTCCCGTGAGCTACCGCATGGCTAAGACTCTCGTGTTTCAGCA
AAGTCAAGACATCCCTTGGCTTGGATCACTGTCACTCTTTTATCAGTGGGACTGCGCCCCTCA
ACCAAGAGACTGCCGAGTTCTTTCTAAGCTTGGACATACCTATAGGCGAGTTGTATGGGTTGA
GTGAGAGCTCGGGACCCACACGATATCCAACCAGAATAACTACAGGCTTCTAAGCTGTGGCA
AGATCTTGACTGGGTGTAAGAATATGCTGTTCCAGCAGAACAAGGATGGCATTGGGGAGATCT
GCCTCTGGGGTAGGCACATCTTCATGGGCTATCTGGAAAGTGAGACTGAACTACAGAGGCCA
TCGATGATGAAGGCTGGCTACACTCTGGGGATCTGGGCCAGCTGGACGGTCTGGGTTTCTCT
ATGTCACCGGCCACATCAAAGAAATCCTTATCACTGCTGGTGGTGAAAATGTGCCCCCATTC
CTGTTGAGACCTTGGTTAAGAAGAAGATCCCATCATCAGTAACGCCATGTTAGTAGGAGATA
AACTGAAGTTTCTGAGCATGTTGCTGACGCTGAAGTGTGAGATGAATCAGATGAGCGGAGAAC
CTCTGGACAAGCTGAACCTTCGAGGCCATCAACTTCTGTCGGGGTCTGGGCAGCCAGGCATCCA
CCGTGACTGAGATTGTGAAGCAGCAAGACCCCTGGTCTACAAGGCCATCCAGCAAGGCATCA
ATGCTGTGAACCAGGAAGCCATGAACAATGCACAGAGGATTGAAAAGTGGGTCACTCTGGAGA
AGGACTTTTCCATCTATGGTGGAGAGCTAGGTCCAATGATGAACTTAAGAGACATTTTGTAG
CCCAGAAATACAAAAACAAATTGATCACATGTACCCTGACTGCTTTGATGGAGCTGCTCTC
AGCTGTTCTGATGCCTTCAGCAGGAAGACCTCATTGCAATAAGTGAAATGCTGCTCTAGGTAG
AAGCTCTCCCTGCTGTTTTTAAGAAGCCACATTCCTCATTGGTCAGTTTCTTGATTGTTGCTC
TGTTGGAGAGGTGCTCCCTAGAAGAACCTGCCATACGTTTCAAAGCAATAAAATCACTGTATA
TCTTTCTAAGGACCTTCAAGTCATGACTCCAGGGAAGCCTATTGGGAAGTCTACTAAAACTG
CCTGATTTACAAGAAAGACCTGAACTTGTGGGCTCCCATTTGATTTTTTCTCCTCAGGGGAC
TCAGACATTAGAAAGAAAAAGCCTCACAGATTTGAAGAACTGGACCCCCAAATCAACTCACCT
GCCTGGAAGCAACTGGGAAACCCTTCCAATAAGTCCTGATAATAAGCACTTCAGGGTCCCAA
AAAAAAAAA

206/550

FIGURE 206

MTIPEFFRESVNRFGTYPALPSKNGKKWEILNFNQYYEACRKAASLIKGLERFHGVGILGF
NSAEWFITAVGAILAGGLCVGIYATNSAEACQYVITHAKVNILLVENDQQQLQKILSIPOSSLE
PLKAI IQYRLPMKKNNNLYSWDDFMELGRSIPDTQLEQVIESQKANQCAVLIYTSGTTGIPKG
VMLSHDNITWIAGAVTKDFKLTDKHETVVSYLPLSHIAAQMMDIWVPIKIGALTYFAQADALK
GTLVSTLKEVKPTVFIGVPQIWEKIHVMVKNSAKSMGLKKKAFVWARNIGFKVNSKMLGKY
NTPVSYRMAKTLVFSKVKTSGLDHCHSFISGTAPLNQETAEFFLSLDIPIGELYGLSESSGP
HTISNQNNYRLLSCGKILTGCKNMLFQQNKDGIGEICLWGRHIFMGYLESETETTEAIDDEGW
LHSGDLGQLDGLGFLYVTGHIKEILITAGGENVPPPIPVETLVKKKIPIISNAMLVGDCLKFLS
MLLTLKCEMNQMSGEPLDKLNFEAINFCRGLGSQASTVTEIVKQQDPLVYKAIQQGINAVNQE
AMNNAQRIEKWVILEKDFSIYGGELGPMMLKRHFVAQKYKKQIDHMYH

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 65-86

N-glycosylation site.

amino acids 196-200

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 282-286

Tyrosine kinase phosphorylation sites.

amino acids 547-555, 608-616

N-myristoylation sites.amino acids 15-21, 74-80, 80-86, 84-90, 185-191, 189-195,
253-259, 337-343, 371-377, 448-454, 536-542**Amidation site.**

amino acids 24-28

Putative AMP-binding domain signature.

amino acids 177-189

Putative AMP-binding domain proteins.

amino acids 173-190

FIGURE 207

[illegible]

208/550

FIGURE 208

MAYRVLGRAGPPQPRRARRLLFAFTLSLSCTYLCYSFLCCDDLGRSRLLGAPRCLRGPSAGG
QKLLQKSRPCDPSGPTPSEPSAPSAPAAVAPAPRLSGSNHSGSPKLGTKRLPQALIVGVKKGG
TRAVLEFIRVHPDVRALGTEPHFFDRNYGRGLDWYRSLMPRTLESQITLEKTPSYFVTQEAPR
RIFNMSRDTKLIVVVRNPVTRAISDYQTLSKKPDIPTEGLSFRNRTLGLVDVSWNAIRIGM
YVLHLESWLQYFPLAQIHVVSGERLITDPAGEMGRVQDFLGIKRFITDKHFYFNKTKGFCLK
KTESSLLPRCLGKSKGRTHVQIDPEVIDQLREFYRPNYKIFYETVGDQFRWE

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 102-106, 193-197, 235-239, 306-310

Tyrosine kinase phosphorylation site.

amino acids 296-305

N-myristoylation sites.

amino acids 51-57, 100-106, 121-127, 125-131

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 20-31

209/550

FIGURE 209

CTTTCCTTATCTGTGTGTACTCTTATCTCACTGTTCTATTTTTTCTCCTCATTTATATTA
CTTTCCTTACCTTTTTTCTGAACTTCTAGGCCTTCTTTTCCAGAACTGGTGGAAGACAAATG
AAACGGCCAAGATGGTAAGAAACAAGCCGCATTTCTCCTTGGGGAGACTGATAATTTAAAAGG
TTTGTGTGTGCAGAAACATTCCCAGCTTCATCACCAACCCTTTCCTTCCACCTCTGCCCCACTG
GAGACCACTTACATCCCGAAGCGGACGCGGCAGCTGAAGTCAGGAAACCATGCATCACATTAG
CAGGAGCCAACTGCAGACTTTAAACTCCGTTCAACATGTGGATGCGGCAGAGAAATGACCTGT
CCAGACAAGCCGGGGCAGCTCATAACTGGTTCATCTGCTCCCTGTGCGTCCCGCGGGTGCGT
AAGCTCTGGAGCAGCCGGCGTCCAAGGACCCGGAGAAACCTTCTGCTGGGCACTGCGTGTGCC
ATCTACTTGGGCTTCTTGGTGTAGCCAGGTGGGGAGGGCCTCTCTCCAGCATGGACAGGCGGT
GAGAAGGGGCCACATCGCAGCCGCGACACCGCCGAGCCATCCTTCCCTGAGATACCCCTGGAT
GGTACCCTGGCCCCCTCCAGAGTCCCAGGGCAATGGGTCCACTCTGCAGCCCAATGTGGTGTAC
ATTACCCTACGCTCCAAGCGCAGCAAGCCGGCCAATATCCGTGGCACCGTGAAGCCCCAAGCGC
AGGAAAAAGCATGCAGTGGCATCGGCTGCCCTAGGGCAGGAGGCTTTGGTTCGGACCATCCCTT
CAGCCGCAGGAAGCGCAAGGGAAGCTGATGCTGTAGCACCTGGGTACGCTCAGGAGCAAAAC
CTGGTTAAGATTGGAGAGCGACCCTGGAGGTTGGTGTGCGGGTCCGGGAGTCCGAGCCGGGGC
CCAGACTTCTGTCAGCCCAGCTCCAGGGAGAGCAACATTAGGATCTACAGCGAGAGCGCCCCC
TCCTGGCTGAGCAAAAGATGACATCCGAAGAATGCGACTCTTGGCGGACAGCGCAGTGGCAGGG
CTCCGGCCTGTGTCTCTAGGAGCGGAGCCCGTTTGTGCTGGTGTGAGGGGGGCGACCTGGC
GCTGTGCTCCGCTGTGGCCCTAGCCCCCTGTGGGCTTCTCAAGCAGCCCTTGGACATGAGTGAG
GTGTTTGCCTTCCACCTAGACAGGATCCTGGGGCTCAACAGGACCCTGCCGTCTGTGAGCAGG
AAAGCAGAGTTCATCCAAGATGGCCGCCCATGCCCCATCATCTTTGGGATGCATCTTTATCT
TCAGCAAGTAATGACACCCATTCTTCTGTAAAGCTCACCTGGGGAACCTATCAGCAGTTGCTG
AAACAGAAATGCTGGCAGAATGGCCGAGTACCCAAGCCTGAATCAGGTTGTACTGAAATACAT
CATCATGAGTGGTCCAAGATGGCACTCTTTGATTTTTTGTACAGATTTATAATCGCTTAGAT
ACAAATTGCTGTGGATTGAGACCTCGCAAGGAAGATGCCTGTGTACAGAATGGATTGAGGCCA
AAATGTGATGACCAAGGTTCTGCGGCTCTAGCACACATTATCCAGCGAAAGCATGACCCAAGG
CATTTGGTTTTTATAGACAACAAGGGTTCTTTGACAGGAGTGAAGATAACTTAAACTTCAAA
TTGTTAGAAGGCATCAAAGAGTTTCCAGCTTCTGCAGTTTCTGTTTTGAAGAGCCAGCACTTA
CGGCAGAAACTTCTTCAGTCTCTGTTTCTTGATAAAGTGTATTGGGAAAGTCAAGGAGGTAGA
CAAGGAATTGAAAAGCTTATCGATGTAATAGAACACAGAGCCAAAATTCTTATCACCTATATC
AATGCACACGGGGTCAAAGTATTACCTATGAATGAATGACAAAAGAATCTTCTGGCTAGGGTG
TTAGATATATTTATGCATTTTTGGTTTTGTTTTTAAATCAAGCACATCAACCTCAAGCCCGTT
TAGCAATGAGGCAGTGTAGATGAATACGTAAAATAAATGACTTTAACCAAGTAGCTATAAAGG
GACTTAGCACTGTATGCATACTTAAAAAGGTTTTGAAAAACAACTACTTGAGAAATATTTGT
TTATATTTTCTCTAACATCATGCTATGTGTGCTGCTGAACATCTGACAACAGAAATTTAGT
TATTATTCTAGCTAAGTTTTGAAAACATTTGTCATGCTGTTAATAGAAAACCTGCAAACCAGA
GATACTGACTCCATTAATAAACCATATTTTGTGCCGTTTTGACTGTTCTGACCAATACTAAT
GGGAACAATTCTTGACGTTTTTCTGTTGCTGATTGTTAACATAGAGCAGTCTCTACACTACCC
TGAGGCAACTCTACATTGGAACACTGAGGCTTACAGCCTGCAAGAGCATCAGAGCTGACCATA
CATTTAAACAGAAATGCTGTTTATTTGCAAAATCACCAGTATATTTTCTATTGTGTCTATAA
AAAATCAGTCATTAAAGTACAAGAATCATATTTTCCATTCTTTTGAAGAAATTTATTTTGTG
TCCCTATGGAAATCATTCACATCTGACAATTTATATGTTAAAGAGTTTTACTCTCTCTATTTT
GGTCCAATTTGTATCTAGTGGCTGAGAAATTAATAATTCTAAAGTATGAAGTTACCTTACTG
AAAATGTACTTACAGAGTATCATTTTTAAATGGATGTCTCTTTAAAAATTTTGTACTTTTAC
CAACAATGTAATATAATTTATGTATATTTTATTAATAATAGTGAATTCCTTAAAAATTTGTCT
ATGTACTTATATTTAATTTGATTTAATGGTTACTGCCAGATATTGAGAAATGGTTCAAATAT
TGAGTGTGTTTCAATAA

210/550

FIGURE 210

MTCPDKPGQLINWFICSLCVPRVRKLWSSRRPRTRRNLLLGTA
CAIYLGFLVSQVGRASLQHG
QAAEKGPHRSDTAEPSFPEIPLDGT
LAPPESQGNSTLQPNVVYITLSKRSK
PANIRGTVK
PKRRKKHAVASAAPGQEALVGPSLQ
PQEAAREADAVAPGYAQGANLVKIG
ERPWLVRGPGVR
AGGPDFLQPSRESNIRIYSESAPSWLS
KDDIRRMRL LADSAVAGLRPVSSR
SGARLLVLEGG
APGAVLRCGPSPCGLLKQPLDMSEV
FAFHLDRI LGLNRTLPSVSRKAEFI
QDGRPCPIILWDA
SLSSASNDTHSSVKLTWGTYQQLLKQ
KCWQNGRVKPKPESGCTEIHHEWSK
MALFDFLLQIYN
RLDTNCCGFRPRKEDACVQNGLRPKC
DDQGSAA LAHIIQRKHDPRHLVFID
NKGFFDRSEDNL
NFKLLEGIKEFPASAVSVLKSQH
LRQKLLQSLFLDKVYWESQGG
RQGI EKLI DVIEHRAKILI
TYINAHGVKVLPMNE

Transmembrane domain:

amino acids 40-56

N-glycosylation sites.

amino acids 98-102, 289-293, 322-326

N-myristoylation sites.amino acids 8-14, 41-47, 97-103, 187-193, 251-257, 252-258,
287-293, 484-490

211/550

FIGURE 211

GTGGGGTGGTGAGCGCAGCGCCGAGGATGAGGAGGTGCAACAGCGGCTCCGGGCCGCCGCCGCTCGCTGCTGCTGC
TGCTGCTGTGGCTGCTCGCGGTTCCCGCGCTAACGCGGCCCGCGGTGCGCGCTCTATTGCGCTTCCGACCCGC
TGACGCTGCTGCAGGCGGACACGGTGCGCGGCGCGGTGCTGGGCTCCCGCAGCGCTGGGCCGTGGAGTTCTTCG
CCTCCTGGTGCGGCCACTGCATCGCCTTCGCCCCGACGTGGAAGGCGCTGGCCGAAGACGTCAAAGCCTGGAGGC
CGGCCCTGTATCTCGCCGCCCTGGACTGTGCTGAGGAGACCAACAGTGCAGTCTGCAGAGACTTCAACATCCCTG
GCTTCCCGACTGTGAGGTTCTTCAAGGCCTTTACCAAGAACGGCTCGGGAGCAGTATTTCCAGTGGCTGGTGTG
ACGTGCAGACGCTGCGGGAGAGGCTCATTGACGCCCTGGAGTCCCATCATGACACGTGGCCCCAGCCTGTCCCC
CACTGGAGCCTGCCAAGCTGGAGGAGATTGATGGATTCTTTGCGAGAAATAACGAAGAGTACCTGGCTCTGATCT
TTGAAAAGGGAGGCTCCTACCTGGGTAGAGAGGTGGCTCTGGACCTGTCCAGCACAAAGGCGTGGCGGTGCGCA
GGGTGCTGAACACAGAGGCCAATGTGGTGAGAAAGTTGGTGTCACCGACTTCCCTCTTGCTACCTGCTGTTC
GGAATGGCTCTGTCTCCCGAGTCCCGTGCTCATGGAATCCAGGTCTTCTATACCGCTTACCTGCAGAGACTCT
CTGGGCTCACCAGGGAGGCTGCCAGACACAGTTGCACCAACCACTGCTAACAAGATAGTCCCCTGTTTGG
AATTGGCAGATCGCTCCAAGATCTACATGGCTGACCTGGAATCTGCACTGCACTACATCCTGCGGATAGAAGTGG
GCAGGTTCCCGGTCTGGAAGGGCAGCGCTGGTGGCCCTGAAAAAGTTTGTGGCAGTGTGGCCAAGTATTTCC
CTGGCCGGCCCTTAGTCCAGAACTTCTGCACTCCGTGAATGAATGGCTCAAGAGGCAGAAAGAAATAAAATTC
CCTACAGTTTCTTTAAACTGCCCTGGACGACAGGAAAGAGGGTGCCGTTCTTGCCAAGAGGTGAACTGGATTG
GCTGCCAGGGAGTGAGCCGCATTTCGGGGCTTTCCCTGCTCCCTGTGGGTCTCTTCCACTTCTTGACTGTGC
AGGCAGCTCGGCAAAATGTAGACCACTCACAGGAAGCAGCCAAGGCCAAGGAGGTCTCCAGCCATCCGAGGCT
ACGTGCACTACTTCTTCGGCTGCCGAGACTGCGCTAGCCACTTCGAGCAGATGGCTGCTGCCTCCATGCACCGGG
TGGGGAGTCCAACGCCGCTGTCTCTGGCTCTGGTCTAGCCACAACAGGGTCAATGCTCGCCTTGCAGGTGCC
CCAGCGAGGACCCCAAGTTCCCAAGGTGCAGTGGCCACCCCGTGAACCTTTGTTCTGCCTGCCACAATGAACGCC
TGGATGTGCCCGTGTGGGACGTGGAAGCCACCTCAACTTCTCAAGGCCCACTTCTCCCAAGCAACATCATCC
TGGACTTCCCTGCAGCTGGGTGAGCTGCCCCGGAGGGATGTGCAGAAATGTGGCAGCCGCCAGAGCTGGCGATGG
GAGCCCTGGAGCTGGAAGCCGGAATTCAACTCTGGACCTGGGAAGCCTGAGATGATGAAGTCCCCACAAACA
CCACCCACATGTGCCGGCTGAGGGACCTGAGGCAAGTGCACCCCGAAGCTGCACCTGGCCTCAGAGCTGCAC
CAGGCCAGGAGCCTCCTGAGCAGATGGCAGAGCTTCAGAGGAATGAGCAGGAGCAGCCGCTTGGGCAGTGGCACT
TGAGCAAGCGAGACACAGGGGCTGCATTGCTGGCTGAGTCCAGGGCTGAGAAGAACCCTCTGGGGCCCTTTGG
AGGTGAGGCGCGTGGGCCGAGCTCCAAGCAGCTGGTGCACATCCCTGAGGGCCAGCTGGAGGCCGAGCTGGAC
GGGGCCGAGGCCAGTGGCTGCAGGTGCTGGGAGGGGGCTTCTTACCTGGACATCAGCCTCTGTGTGGGGCTCT
ATTCCCTGTCTTTCATGGGCCCTGTGGCCATGTACACCTACTTCCAGGCCAAGATAAGGGCCCTGAAGGGCCATG
CTGGCCACCCTGCAGCTGAACCACTGGGGAGGAGGCGGGAGAGGGAGCTGCCATCTCTAGGCACCTCAAGCCC
CCTGACCCCATTCCTCCCTCCCACCCCTTGCTCCTTGCTGGCCTAGAAGTGTGGGAAATTGAGGAAAACGAG
TTGCTCCAGTGAAGCTTCTTGGGGTTGCTAGGACAGAGAGCTCCTTTGACACAAAAGACAGGAGCAGGGTCCAGG
TTCCCTGCTGTGCAGGGAGGGCAGCCCCGGGCGAGTGGGCATAGGGCAGCTCAGTCCCTGGCCTCTTAGCACCAC
ATTCTGTTTTTTCAGCTTATTTGAAGTCTGCTCATTCTCACTGGAGCCTCAGTCTCTCTGCTTGGTCTTGGC
CCTCAACTGGGGCAAGTGAAGCCAGAGGAGGGTCCCCAGCTGGGTGGGCTGGAATGGAATCCTCACTAGCTGC
TGGGGCTCCGCCACCCCTGCTCCCTTCCGGACAATGAAGAAGCCTTTGCACCTGGGAGGAAGGACCACCCCGG
CCCTCTATGCTGGCCAGCCTCCAGCTCCTCAGACCTCCTGGGTGGGGTTTGGCTTCAAGGTGGGGTTTGAAGC
TTCTGGAAGTCGTGCTGGTCTCCAGGTGAGGCAAGCCATGGTTGCTGGGCTGTAGGCTGAGTGGCTTGGTGGT
GGGACCTGACGAGTTGGTGGCATGGGAAGGATGTGGGTCTCTAGTGCCTTGCCCTGGCTTAGCTGCAGGAGAAGA
TGGCTGCTTTCACTTCCCCCATTTAGCTCTGCTCCCTCTGAGCCTGGTCTTTTGTCTTTTATTTTGGTCTC
CAAGATGAATGCTCATCTTTGGAGGTGCCAGGTAGAAGCTAGGGAGGGGAGTGTCTCTCTCCAGGTTTCAC
CTTCCAGTGTGCAGAAGTTAGAAGGGTCTGGCGGGGGCAGTGCCTTACACATGCTTGATTCCCACGCTACCCCT
GCCTTGGGAGGTGTGTGGAATAAATTATTTTGTTAAGGCA

212/550

FIGURE 212

MRRCSGSGPPPSLLLLLLWLLAVPGANAAPRSALYSPSDPLTLLQADTVRGAVLGSRSAWAV
EFFASWCGHCIAFAPTWKALAEDVKAWRPALYLAALDCAETNSAVCRDFNIPGFPTVRFFKA
FTKNGSGAVFPVAGADVQTLRERLIDALESHHDTWPPACPPLEPAKLEEIDGFFARNNEEYLA
LIFEKGGSYLGREVALDLSQHKGVAVRRVLNTEANVVRKFGVTDFFPCYLLFRNGSVSRVPVL
MESRSFYTAYLQRLSGLTREAAQTTVAPTTANKIAPT VWKLADRSKIYMADLESALHYILRIE
VGRFPVLEGQRLVALKKFVAVLAKYFPGRPLVQNFLHSVNEWLKRQKRNKI PYSFFKTALDDR
KEGAVLAKKVNWIGCQGSEPHFRGFPCSLWVLFHFLT VQAARQNV DHSQEAAKAKEVLPAIRG
YVHYFFGCRDCASHFEQMAAASMRVGGSPNAAVLWLWSSSHNRVNARLAGAPSEDPQFPKVQWP
PRELCSACHNERLDVPVWDVEATLNLFLKAHFSPSNIILDFFPAAGSAARRDVQNVAAPPELAMG
ALELESRNSTLDPGKPEMMKSPTNTTPHVPAEGPEASRPPKLHPGLRAAPGQEPPEHMAELQR
NEQEQPLGQWHLSKRD TGAALLAESRAEKNRLWGP LEVRRVGRSSKQLVDIPEGQLEARAGRG
RGQWLQVLGGGFSYLDISLCVGLYSLSFMGLLAMYT YFQAKIRALKGHAGHPAA

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 705-728

N-glycosylation sites.

amino acids 130-134, 243-247, 575-579

Glycosaminoglycan attachment site.

amino acids 6-10

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 644-648

N-myristoylation sites.amino acids 52-58, 56-62, 196-202, 381-387, 392-398, 448-454,
468-474, 684-690, 702-708**Cytochrome c family heme-binding site signature.**

amino acids 509-515

Thioredoxin family proteins

amino acids 62-78

213/550

FIGURE 213

GCACGAGGCCGACTTCCAGACCATCTACAACCTGCACGGCCTGGAACAGCTTCGGCTCCGACAC
TGAGATCATCCGGCTCAAGGAGCAAGGTTTCGGAATGAAGTCGGGAGCCGGGCTGGAAGCAGA
GTCTGTGCCG**ATG**GCCGTCATCATTGGGGTGGCCGTAGGAGCTGGTGTGGCCTTCCTCGTCCT
TATGGCAACCATCGTGGCGTTCTGCTGTGCCCGTTCCCAGAGAAATCTCAAAGGTGTTGTGTC
AGCCAAAATGATATCCGAGTGGAATTTGTCCACAAGGAACCAGCCTCTGGTCGGGAGGGTGA
GGAGCACTCCACCATCAAGCAGCTGATGATGGACCGGGTGAATTCAGCAAGACTCAGTCCT
GAAACAGCTGGAGGTCCTCAAAGAAGAGGAGAAAGAGTTTCAGAACCTGAAGGACCCACCAA
TGGCTACTACAGCGTCAACACCTTCAAAGAGCACCACTCAACCCCGACCATCTCCCTCTCCAG
CTGCCAGCCGACCTGCGTCCTGCGGGTAAGCAGCGTGTGCCACAGGCATGTCTTCACCAA
CATCTACAGCACCTGAGCGGCCAGGGCCGCCTCTACGACTACGGGCAGCGGTTTGTGCTGGG
CATGGGCAGCTCGTCCATCGAGCTTTGTGAGCGGGAGTTCCAGAGAGGCTCCCTCAGCGACAG
CAGTCCTTCCTGGACACGCAGTGTGACAGCAGCGTCAGCAGCAGCGGCAAGCAGGATGGCTA
TGTGCAGTTCGACAAGGCCAGCAAGGCTTCTGCTTCCTCCTCCCACCACTCCCAGTCCTCGTC
CCAGAACTCTGACCCAGTCGACCCCTGCAGCGGCGGATGCAGACTCACGTCT**TAAG**GATCACA
CACCGCGGGTGGGGACGGGCCAGGGAAGAGGTCAGGGCACGTTCTGGTTGTCCAGGGACGAGG
GGTACTTTGCAGAGGACACCAGAATTGGCCACTTCCAGGACAGCCTCCCAGCGCCTCTGCCAC
TGCCCTTCCTTGAAGCTCTGATCAAGCACAAATCTGGGTCCCCAGGTGCTGTGTGCCAGAGGT
GGGCGGGTGGGGAGACAGACAGAGGCTGCGGCTGAGTGGCGCTGTGCTTAGTGCTGGACACCCG
TGTCCCCGGCCCTTTCCTGGAGGCCCTCTACCACCTGCTCTGCCACAGGCACAAGTGGCAG
CTATAACTCTGCTTTCATGAAACTGCGGTCCACTCTCTGGTCTCTCTGTGGGCTCTACCCCTC
ACTGACCACAAGCTCTACCTACCCCTGTGCCTGTGCTCCCATAACAGCCCTGGGGAGAAGGGGA
TGACGTCTTCCCAGCACTGAGCTGCCCCAGAAACCCCGGCTCCCCACTGCTGCTCATAGCCCA
TACCCTGGAGGCTGACAAGCCAGAAATGGCCTTGGCTAAAGGAGCCTCTCTCTCACCAGGCTG
GCCGGGAGCCCACCCCAATTTGTTTGGTGTGTTTGTGTCCATACTCTTGCAAGTTCTGTCCTTG
GACTTGATGCCGCTGAACTCTGCGGTGGGACCGGTCCCGTCAGAGCCTGGTGTACTGGGGGGA
GGGAGGGAGGAGGGAGCCTGTGCTGACGGAGCACCTCGCCGGGTGTGCCCTCCTGGGCTGTG
TGACCCAGCCTCCCCACCCACCTCCTGCTTTGTGTACTCCTCCCCCTCCCCCTCAGCACAATC
GGAGTTCATATAAGAAGTGCGGGAGCTTCTCTGGTCAGGGTCTCTGAACACTTATGGAGAGA
GTGCTTCCTGGGAAGTGTGGCGTTTGAAGGGGCTGGAGGGCAGGTCTTTAAGATGGCGAGACT
GCCCTTCTCAGCTGATAAACACAAGAACGGCGATCCTGTCTTCAGTAAGGCTCCACGAGAAGA
GAGGAAGTATATCTACACCTCAACCCTCCTAGTCACCACCTGAAATAAATGTTAGGGAAAAAAA

214/550

FIGURE 214

MAVIIGVAVGAGVAFLVLMATIVAFCCARSQRNLKGVVSAKNDIRVEIVHKEPASGREGEEHS
TIKQLMMDRGEFQQDSVLKQLEVLKEEEKEFQNLKDPTNGYYSVNTFKEHHSTPTISLSSCQP
DLRPAGKQRVPTGMSFTNIYSTLSGQGRLYDYGQRFVLGMGSSSIELCEREFQRGSLSDSSSF
LDTQCDSSVSSSGKQDGYVQFDKASKASASSSHHSQSSSQNSDPSRPLQRRMQTHV

Signal peptide:

amino acids 1-28

Glycosaminoglycan attachment site.

amino acids 150-154

N-myristoylation sites.

amino acids 6-12, 10-16, 36-42, 139-145, 165-171

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 114-125

215/550

FIGURE 215

CAGCCTTCCTCCCCCAGCCTGAGTGACTACTCTATTCCTTGGTCCCTGCTATTGTCTGGGGACG
ATTGCATGGGCTACGCCAGGAAAGTAGGCTGGGTGACCGCAGGCCTGGTGATTGGGGCTGGCG
CCTGCTATTGCATTTATAGACTGACTAGGGGAAGAAAACAGAACAAGGAAAAAATGGCTGAGG
GTGGATCTGGGGATGTGGATGATGCTGGGGACTGTTCTGGGGCCAGGTATAATGACTGGTCTG
ATGATGATGATGACAGCAATGAGAGCAAGAGTATAGTATGGTACCCACCTTGGGCTCGGATTG
GGACTGAAGCTGGAACCAGAGCTAGGGCCAGGGCAAGGGCCAGGGCTACCCGGGCACGTCGGG
CTGTCCAGAAACGGGCTTCCCCAATTCAGATGATACCGTTTTGTCCCCTCAAGAGCTACAAA
AGGTTCTTTGCTTGGTTGAGATGTCTGAAAAGCCTTATATTCTTGAAGCAGCTTTAATTGCTC
TGGGTAACAATGCTGCTTATGCATTTAACAGAGATATTATTCGTGATCTGGGTGGTCTCCCAA
TTGTCTGCAAAGATTCTCAATACTCGGGATCCCATAGTTAAGGAAAAGGCTTTAATTGTCCTGA
ATAACTTGAGTGTGAATGCTGAAAATCAGCGCAGGCTTAAAGTATACATGAATCAAGTGTGTG
ATGACACAATCACTTCTCGCTTGAATCATCTGTGCAGCTTGCTGGACTGAGATTGCTTACAA
ATATGACTGTTACTAATGAGTATCAGCACATGCTTGCTAATTCCATTTCTGACTTTTTTCGTT
TATTTTCAGCGGGAAATGAAGAAACCAAACCTTCAGGTTCTGAAACTCCTTTTGAATTTGGCTG
AAAATCCAGCCATGACTAGGGAACCTGCTCAGGGCCCAAGTACCATCTTCACTGGGCTCCCTCT
TTAATAAGAAGGAGAACAAAGAAGTTATTCTTAAACTTCTGGTCATATTTGAGAACATAAATG
ATAATTTCAAATGGGAAGAAAATGAACCTACTCAGAATCAATTCGGTGAAGGTTCACTTTTTT
TCTTTTTTAAAAGAATTTCAAGTGTGTGCTGATAAGGTTCTGGGAATAGAAAGTCACCATGATT
TTTTGGTGAAAGTAAAAGTTGGAAAATTCATGGCCAAACTTGCTGAACATATGTTCCCAAAGA
GCCAGGAATTAACACCTTGATTTTGTAATTTAGAAGCAACACACATTGTAAACTATTTCATTTTC
TCCACCTTGTTTATATGGTAAAGGAATCCTTTCAGCTGCCAGTTTTGAATAATGAATATCATA
TTGTATCATCAATGCTGATATTTAACTGAGTTGGTCTTTAGGTTTAAAGATGGATAAATGAATA
TCACTACTTGTTCTGAAAACATGTTTGTGCTTTTTATCTCGCTGCCTAGATTGAAATATTTT
GCTATTTCTTCTGCATAAGTGACAGTGAACCAATTCATCATGAGTAAGCTCCCTTCTGTCATT
TTCATTGATTTAATTTGTGTATCATCAATAAAATTGTATGTTAATGCTGGAAAGA

216/550

FIGURE 216

MGYARKVGWVTAGLVIGAGACYCIYRLTRGRKQNKKEKMAEGGSGDVDDAGDCSGARYNDWSDD
DDDSNESKSIVWYPPWARIGTEAGTRRARARARARATRRARRAVQKRASPNSSDDTVLSPQELQKV
LCLVEMSEKPYILEAALIALGNNAAYAFNRDIIRDLGGLPIVAKILNTRDPIVKEKALIVLNN
LSVNAENQRRLKVYMNQVCDDTITSRLNSSVQLAGLRLLTNMTVTNEYQHMLANSISDFFRLF
SAGNEETKLQVLKLLLNLAENPAMTRELLRAQVPSSLGSLFNKKENKEVILKLLVIFENINDN
FKWEENEPTQNQFGEGSLFFFLKEFQVCADKVLGIESHHDFLVKVKVGKFMAKLAEHMFPSQE

Signal peptide:

amino acids 1-20

N-glycosylation sites.

amino acids 68-72, 189-193, 217-221, 230-234

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 107-111

N-myristoylation sites.amino acids 13-19, 17-23, 19-25, 54-60, 83-89, 147-153, 255-261,
290-296**Amidation site.**

amino acids 29-33

217/550

FIGURE 217

[illegible]

218/550

FIGURE 218

MAIAQLATEYVFSDFLLEKEPTEPKFKGLRLELAVDKMVTCTIAVGLPLLLISLAFAQEISIGTQ
ISCFSPSSFSWRQA AFVDSYCWA AVQQKNSLQSESGNLPLWLHKFFPYILLFFAILLYLPPLF
WRFAAAPHCSDLKFIMEELDKVYNRAIKA AKSARDLDMRDGACSVPGVTENLGQSLWEVSES
HFKYPIVEQYLKTKKNSNNLI IKYISCRLLTLIIILLACIYLGYYFSLSSLSDEFVCSIKSGI
LRNDSTVPDQFQCKLIAVGIFQLLSVINLVVYVLLAPVVVYTLFVFPFRQKTDVLKVYEILPTF
DVLHFKSEGYNDLSLYNLFLEENISEVKS YKCLKVLENIKSSGQGIDPMLLLTNLGMIMDVV
DGKTPMSAEMREEQGNQTAE LQGMNIDSETKANNGEKNARQRLLDSSC

Transmembrane domains:

amino acids 37-55, 108-126, 216-232, 273-290

N-glycosylation sites.

amino acids 255-259, 338-342, 394-398

Glycosaminoglycan attachment site.

amino acids 357-361

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 203-207

N-myristoylation sites.

amino acids 61-67, 174-180, 251-257, 393-399

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 218-229

219/550

FIGURE 219

CTGTGAGTGACACACGCTGAGTGGGGTGAAGGGAAATGCTGGTGAATTTTCATTTTGAGGTGTG
GGTTGCTGTTAGTCACTCTGTCTCTTGCCATTGCCAAGCACAAGCAATCTTCCTTCACCAAAA
GTTGTTACCCAAGGGGAACATTGTCCCAAGCTGTTGACGCTCTCTATATCAAAGCAGCATGGC
TCAAAGCAACGATTCCAGAAGACCGCATAAAAAATATACGATTATTAAAAAAGAAAACAAAA
AGCAGTTTATGAAAACTGTCAATTTCAAGAACAGCTTCTGTCCCTTCTTCATGGAAGACGTTT
TTGGTCAACTGCAATTGCAAGGCTGCAAGAAAATACGCTTTGTGGAGGACTTTCATAGCCTTA
GGCAGAAATTGAGCCACTGTATTTCTGTGCTTCATCAGCTAGAGAGATGAAATCCATTACCA
GGATGAAAAGAATATTTTATAGGATTGGAAACAAAGGAATCTACAAAGCCATCAGTGAAGTGG
ATATTCTTCTTTCCTGGATTAAAAAATTATTGGAAAGCAGTCAGTAAACCAAAGCCAAGTACA
TTGATTTTACAGTTATTTTGAAATACAATAAGAACTGCTAGAAATATGTTTATAACAGTCTAT
TTCTTTTAAAACTTTTTAACATAATACTGACGGCATGTTAGGTGATTCAGAATAGACAAGAA
GGATTTAGTAAATTAACGTTTTGGATATAAGTTGTCATAATTTGCACATTTTCTGTGTTTTT
AAATAATGTTTCCATTCTGAACATGTTTTGTCATTCACAAGTACATTGTGTCAACTTAATTTA
AAGTATGTAACCTGAATTAACGTCGTGTAATATTTGTGTGTGGAGTGGGATGTGGGGGGTGGAG
GGGAATGACAGATTTCTGGAATGCAATGTAATGTTACTGAGACTTAAATAGATGTTATGTAT
ATGATTGTCTGTTTAAGTGTTTGAAAATTGTTAATTATGCCCAGTGTGAAGTCTAGTACTTAAC
ACATTTTGATTTTAATTAAATAAATTGGGTTTCCTTCTCAAAAAAAAAAAAAAAAAAAAAA
AAAAA

220/550

FIGURE 220

MLVNFILRCGLLLVTLSLAIKHKQSSFTKSCYPRGTLSQAVDALYIKAAWLKATIPEDRIKN
IRLLKKKTKKQFMKNCQFQEQLLSFFMEDVFGQLQLQGCKKIRFVEDFHSLRQKLSHCISCAS
SAREMKSITRMKRIFYRIGNKGIYKAISELDILLSWIKKLLESSQ

Signal sequence:

amino acids 1-21

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 68-71

N-myristoylation site.

amino acids 148-153

Interleukin-10 proteins.

amino acids 58-94, 74-102, 128-170

221/550

FIGURE 221

GACCACGGCCCTGCGCCCCAGCCAGGCCTGAGGACATGAGGCGGCCGGCGGGCGGTGCCGCTCC
TGCTGCTGCTGTGTTTTGGGTCTCAGAGGGCCAAGGCAGCAACAGCCTGTGGTCGCCCCAGGA
TGCTGAACCGAATGGTGGGCGGGCAGGACACGCAGGAGGGCGAGTGGCCCTGGCAAGTCAGCA
TCCAGCGCAACGGAAGCCACTTCTGCGGGGGCAGCCTCATCGCGGAGCAGTGGGTCCTGACGG
CTGCGCACTGCTTCCGCAACACCTCTGAGACGTCCCTGTACCAGGTCCTGCTGGGGGCAAGGC
AGCTAGTGCAGCCGGGACCACACGCTATGTATGCCCGGGTGAGGCAGGTGGAGAGCAACCCCC
TGTACCAGGGCACGGCCTCCAGCGCTGACGTGGCCCTGGTGGAGCTGGAGGCACCAAGTGCCTT
TCACCAATTACATCCTCCCCGTGTGCCTGCCTGACCCCTCGGTGATCTTTGAGACGGGCATGA
ACTGCTGGGTCACTGGCTGGGGCAGCCCCAGTGAGGAAGACCTCCTGCCCCGAACCGCGGATCC
TGCAGAAACTCGCTGTGCCCATCATCGACACACCCAAGTGCAACCTGCTCTACAGCAAAGACA
CCGAGTTTGGCTACCAACCCAAAACCATCAAGAATGACATGCTGTGCGCCGGCTTCGAGGAGG
GCAAGAAGGATGCCTGCAAGGGCGACTCGGGCGGCCCCCTGGTGTGCCTCGTGGGTGAGTCGT
GGCTGCAGGCGGGGGTGATCAGCTGGGGTGAGGGCTGTGCCCGCCAGAACCGCCCAGGTGTCT
ACATCCGTGTCACCGCCCACCACAACCTGGATCCATCGGATCATCCCCAACTGCAGTTCACGC
CAGCGAGGTTGGGCGGCCAGAAGTGAGACCCCCGGGGCCAGGAGCCCCCTTGAGCAGAGCTCTG
CAGGAGCCTGCCCCGCCACACCATCCTGCTGGTCCTCCCAGCGCTGCTGTTGCACCTGTGAG
CCCCACCAGACTCATTTGTAAATAGCGCTCCTTCCCTCTCAAATACCCTTATTTTATTT
ATGTTTCTCCCAATAAAAACCCAGCCTGTGTGCCAGCTGAAAAAAAAAAAAAAAAAAAAA

222/550

FIGURE 222

MRRPAAVPLLLLLCFGSQRAKAATACGRPRMLNRMVGGQDTQEGEWPWQVSIQRNGSHFCGGS
LIAEQWVLTAAHCFRNTSETSLYQVLLGARQLVQPGPHAMYARVRQVESNPLYQGTASSADVA
LVELEAPVPFTNYILPVCLPDPSVIFETGMNCWVTGWGSPSEEDLLPEPRILQKLAVPIIDTP
KCNLLYSKDTEFGYQPKTIKNDMLCAGFEEGKKDACKGDSGGPLVCLVGQSWLQAGVISWGEG
CARQNRPGVYIRVTAHHNWIHRIIPKLQFQPARLGGQK

Important features of the protein:**Signal peptide:**

amino acids 1-22

N-glycosylation sites.

amino acids 55-58, 79-82

Casein kinase II phosphorylation sites.

amino acids 121-124, 165-168, 167-170, 248-251

Tyrosine kinase phosphorylation sites.

amino acids 78-86, 197-203

N-myristoylation sites.

amino acids 16-21, 37-42, 56-61, 62-67, 118-123

Amidation site.

amino acids 219-222

Serine proteases, trypsin family, histidine active site.

amino acids 71-76

223/550

FIGURE 223

CAAGATGTGGACAGCTCTTGTGCTCATTTGGATTTTCTCCTTGTCTTATCTGAAAGCCATGC
GGCATCCAACGATCCACGCAACTTTGTCCCTAACAAAATGTGGAAGGGATTAGTCAAGAGGAA
TGCATCTGTGGAACAGTTGATAATAAAACGTCTGAGGATGTAACCATGGCAGCAGCTTCTCC
TGTACATTGACCAAAGGGACTTCGGCAGCCCACCTCAACTCTATGGAAGTCACAACAGAGGA
CACAAGCAGGACAGATGTGAGTGAACCAGCAACTTCAGGAGTTGCAGCTGATGGTGTGACCTC
CATTGCTCCACGGCTGTGGCCTCCAGTACGACTGCGGCCTCCATTACGACTGCGGCCTCCAG
TATGACTGTGGCCTCCAGTGCTCCACGACTGCAGCCTCCAGTACAACTGTGGCCTCCATTGC
TCCCACGACTGCAGCCTCCAGTATGACTGCGGCCTCCAGCACTCCCATGACACTTGCCTCCC
CGCGCCACGTCCACTTCCACAGGGCGGACCCCGTCCACTACCGCCACTGGGCATCCATCTCT
CAGCACAGCCCTCGCACAAGTGCCAAAGAGCAGCGGTTGCCAAGAACAGCAACCCTGGCCAC
ATTGGCCACACGTGCTCAGACTGTAGCGACCACAGCAAACACAAGCAGCCCCATGAGCACTCG
TCCAAGTCCTTCCAAGCACATGCCCAGTGACACCGCGGCAAGCCCTGTACCCCTATGCGTCC
CCAAGCACAAGGTCCCATTAGCCAGGTGTGAGTGGACCAGCCTGTGGTTAACACAACAAATAA
ATCCACACCCATGCCCTCAAACACAACCCAGAGCCCGCCCCACCCACAGTGGTGACCAC
CACCAAGGCACAAGCCAGGGAGCCAACTGCCAGCCAGTGCCAGTACCTCACACCAGCCCAAT
CCCTGAGATGGAGGCCATGTCCCCACGACACAGCCAAGCCCCATGCCATATACCCAGAGGGC
CGCTGGGCCAGGCACATCCCAGGCACCGGAGCAGGTAGAGACTGAAGCCACACCAGGTACTGA
TTCCACTGGGCCAACACCCAGGAGCTCAGGGGGCACTAAGATGCCAGCCACGGACTCGTGCCA
GCCCAGCACCCAAGGCCAGTACATGGTGGTGCTGTTACTCGGGGTGACCCCTTTTCATCACAGTCTTGGTTTT
GTTTGCCCTGCAGGCCTATGAGAGCTACAAGAAGAAGGACTACACCCAGGTGGACTACTTAAT
CAACGGGATGTATGCGGACTCAGAAATGTGAGGGGGGCGGGGGCCTGGCGGGAGGCCTGGCCC
CTTCCTCGTCCTTTTCCTTTTGCTTTGAGACCAAACCAAGTGCTTCCAAATTCTTTTGGTGCA
ATTGAGGAGATATGCCAGATGCTTAAACACATTTAATTGCTGTCAGATTAATTCCATGATCAC
TAAAGAGTTGCTGCTTTTTTCATATTTATTTTGTAAATGATTCTGTGCCAGGAGCAGCTGG
GGGTTCCACCTCAGGGTGGGGCGGCAGGACCCCGTCTCCCAGGTGTCGGAGCCTGACCTGA
ATTAAAGTACTGACTGCTCGCCA

224/550

FIGURE 224

MWTALVLIWIFSLSLSESHAASNDPRNFVPNKMWKGLVKRNASVETVDNKTSEDVTMAAASPV
TLTKGTSAAHLNSMEVTTEDTSRTDVSEPATSGVAADGVTSIAPTAVASSTTAASITTAASSM
TVASSAPTTAASSTTVASIAPTTAASSMTAASSTPMTLALPAPTSTSTGRTPSTTATGHPSLS
TALAQVPKSSALPRTATLATLATRAQTVAATTANTSSPMSTRPSPSKHMPSDTAASPVPPMRPQ
AQGPISQVSVDQPVVNTTNKSTPMPSNTTPEPAPTPTVVTTTKAQAREPTASPVVPHTSPIP
EMEAMSPTTQPSMPYTQRAAGPGTSQAPEQVETEATPGTDSTGPTPRSSGGTKMPATDSCQP
STQGQYMVVTTEPLTQAVVDKTL LLVLLLGVTLFITVLVLFALQAYESYKKKDYTQVDYLIN
GMYADSEM

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 396-420

N-glycosylation sites.

amino acids 41-44, 49-52, 222-225, 268-271, 271-274

Casein kinase II phosphorylation sites.

amino acids 14-17, 51-54, 80-83, 85-88, 280-283, 434-437

N-myristoylation sites.

amino acids 68-73, 354-359

Aldo/keto reductase family putative active site signature.

amino acids 195-210

225/550

FIGURE 225

GGAAGGCGCTCAAGGTGCGCGGCCCGGGGCGCGCTACTGGGGGCGCCCTCCGCGGTGGGCAGC
GCGCCAGGGATCGGCCTGGGCAGCCGCGGGGCGCGCGAAGGCTGCGCTTTCCTACGGCCCCC
CTCGCTTCCTCCGGCACGGCGGCAACGGAGATTTCCTCTCGGGGAACTACGCGGATCCTTTT
CGGGGATCCTCGCCCCGCCCGAGTTCTCCGCCCCCTCCCCTTTGCTGGGGCGCCTGGGCTGGC
CCGCGCAGGGGAGGAGGCTCTGGCAGCCTGGGCAGGGAGGCGGCGGGGGGCCGCGGAGCCGCT
GGCCATCGATTCTCCCCGCCATGTGACGCCGTCTTAGCCCTGCGACCCCCAGCGCTCCCGG
GCCTGCGCCTCCGCCCCGCCGCGCAGCGCACG**ATG**CCTTCTGCCGGGACGCGCACGCCAACC GC
CGACGCCCCAGCCCGTGACGATCCCGGCCCTCCGCCGCGAGGTAGAGCCGCCGGGGCAGCTCC
TGCGCCTCTTCTACTGCACTGTCTGGTCTGCTCCAAAGAGATCTCAGCGCTCACCGACTTCT
CTGGTTACCTAACCAAACTCCTGCAAAACCACACCACCTATGCCTGTGATGGGGACTATTTGA
ATCTACAGTGCCCTCGGCATTCTACGATAAGTGTCCAATCGGCATTTTATGGGCAAGATTACC
AAATGTGTAGTTCCCAGAAGCCTGCCTCCCAGAGGGAAGACAGCTTAACCTGTGTGGCAGCCA
CCACCTTCCAGAAGGTGCTGGACGAATGCCAGAACCAGCGGGCCTGCCACCTCCTGGTCAATA
GCCGTGTTTTTGGACCTGACCTTTGTCCAGGAAGCAGTAAATACCTCCTGGTCTCCTTTAAAT
GCCAACCTAATGAATTAAAAACAAAACCGTGTGTGAAGACCAGGAGCTGAACTGCACTGCC
ATGAATCCAAGTTCTCAACATCTACTCTGCGACCTACGGCAGGAGGACCCAGGAAAGGGACA
TCTGCTCCTCCAAGGCAGAGCGGCTCCCCCTTTTCGATTGCTTGTCTTACTCAGCTTTGCAAG
TCCTATCCCGAAGGTGCTATGGGAAGCAGAGATGCAAAATCATCGTCAACAATCACCATTTTG
GAAGCCCCTGTTTGCCAGGCGTGAAAAAATACCTCACTGTGACCTACGCATGTGTTCCCAAGA
ACATACTCACAGCGATTGATCCAGCCATTGCTAATCTAAAACCTTCTTTGAAGCAGAAAGATG
GTGAATATGGTATAAACTTCGACCCAAGCGGATCGAAGGTTCTGAGGAAAGATGGAATTCTTG
TTAGCAACTCTCTGGCAGCCTTTGCTTACATTAGAGCCCACCCAGAGAGAGCTGCCCTGCTGT
TCGTGTCCAGTGTCTGCATCGGCCTGGCCCTCACACTGTGCGCCCTGGTCATCAGAGAGTCCT
GTGCCAAGGACTTCCGCGACTTGACGCTGGGGAGGGAGCAGCTGGTGCCAGGAAGTGACAAGG
TCGAGGAGGACAGCGAGGATGAAGAAGAGGAGGAGGACCCCTCTGAGTCTGATTTCCAGGGG
AACTGTGCGGGGTTCTGTAGGACTTCATATCCTATATACAGTTCCATAGAAGCTGCAGAGCTCG
CAGAAAGGATTGAGCGCAGGGAGCAAATCATTAGGAAATATGGATGAACAGTGGTTTGGACA
CCTCGCTCCCAAGAAACATGGGCCAGTTCTACT**TGA**AAACCACATGCATCTTGATGCGATCGCA
CTTTCTGAAGAAGGAAGGATCCCAAATGCCCTCCAGTTCTGGTTACCTGTACCTTCTATGA
AGGAGAATTCGTCACTGTCATTCAACACTCGTGAGGCCAGGAAGCTATTAAAGGGATGTTCAA
GCTGTTTCTAGCACATTCCAAAATAAATGAGGAGGGAGGAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

226/550

FIGURE 226

MLLPGRARQPPTPQPVQHPGLRRQVEPPGQLRLFYCTVLVCSKEISALTDGSGYLTKLLQNH
TTYACDGDYLNLCQPRHSTISVQSAFYGQDYQMCSSQKPAQREDSLTCVAATTFQKVLDECQ
NQRACHLLVNSRVFGPDLCPGSSKYLVSFKCQPNELKNKTVCEDQELKLHCHESKFLNIYSA
TYGRRTQERDICSSKAERLPPFDCLSYALQVLSRRCYGKQRCKIIVNNHHFGSPCLPGVKKY
LTVTYACVPKNILTAIDPAIANLKPSLKQKDGEYGINFDPGSKVLRKDGILVSNLSAAFAYI
RAHPERAALLFVSSVCIGLALTLCALVIRESCAKDFRDLQLGREQLVPGSDKVEEDSEDEEEE
EDPSESDFPGELSGFCRTSYPIYSSIEAAELAERIERREQIIQEIMNSGLDTS�PRNMGQFY

Transmembrane domains:

amino acids 32-49, 322-343

N-glycosylation sites.

amino acids 62-66, 165-169

Tyrosine kinase phosphorylation site.

amino acids 280-287

N-myristoylation site.

amino acids 302-308, 333-339, 428-434

Amidation site.

amino acids 191-195

227/550

FIGURE 227

GGCACGAGGTGGAAGGGCTTTTACAAACAGATTGCTGGCCCCACCCCCAGAATTTCTCATCA
GGAGTGGGCAAGACCAATCATTTCGATTTCTGACAAGTTCCCAGGAGCTGCAGCTGCTGGCCC
TGGAACCACACTTTGAGAACCACTGCTTTAGACCAAACACCAAAGGAAGATGCAGCCACCCTC
CTTTACATGTCACAACGCTCAGGGTCCATGAGTACCTCAGGCTGTCCAGCTGAGCTCCACCTG
CAGCAGCCGAGATTCCCGACTCGCTCCACCATTGGGGGCTAGGAGTGAAGCGTGTCAACCATGG
TCAGCTCATGGCCAGCCAGGAAAGCCTCTCTGCTGTGCGTCTGTGCAGTTCTTGTTCTTCCCT
GGAGGACTCTTGATCGCCTGTGATCTTGGCCAGGAGACCAGGTGCCTGGGTCCCTTCCTGGA
AGGGGACAAGTTACACACCCCAGCCCCATTTTCCCACCAACTTCTACATGCCTTGGGAGAACC
TTCTACATGTTGGCTGCCCCCTTCCCCTATTTTCAGCAGTGCCCAGTCCTGCTTATAAACCTGA
GGCCTGCTCCCCATACCTTCCCTGTGCAAGTGCCAGCCGTTATTCCAGGCAGCCCAATGTTGT
TGAGGCCAGATGGATTCCCTGGAAGCAGCTGGCCCATGGATGTGAGTCATCACAGTATTCTAGA
AACAGAGAAGAGGTCTTAACCTAATGCGCATAGAGAAATTGTTCTCATTGTAAACATACCCCT
GTCCTTAGCTGATCTAGGTGGAAGCCCAGCTTCATGTGCTAGGGGGCATGATAATGATAATAA
AGGAATTGTATCTAGGACTAA

228/550

FIGURE 228

MVSSWPARKASLLCVC AVLVL PWRTL GSPVILARRPGAWVPSWKGTSYTPQPHEPTNFYMPWE
NLLHVGCP LPLFQQCPVLLINLRPAPHTFPVQVPAVIPGSPMLLRPDGFLEAAGPWM

Signal peptide:

amino acids 1-27

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 8-12

229/550

FIGURE 229

GGGAAGGGATGCAAGGAAGCCCTCCGGCGCTGCGCTCCGAGGCGGGAGACAGCGTCCCGCTGA
AAATGTGTGTCTGACATGCAAGCTCAGTGGGGCAGAGACCCGTGGATTGCTGTGCCCTGCCCT
CCGGACCTGGATC**ATGA**AGGTGTTGGGAAGAAGCTTCTTCTGGGTGCTGTTTCCCGTCCTTCC
CTGGGCGGTGCAGGCTGTGGAGCACGAGGAGGTGGCGCAGCGTGTGATCAAACCTGCACCGCGG
GCGAGGGGTGGCTGCCATGCAGAGCCGGCAGTGGGTCCGGGACAGCTGCAGGAAGCTCTCAGG
GCTTCTCCGCCAGAAGAATGCAGTTCTGAACAACTGAAAACCTGCAATTGGAGCAGTGGAGAA
AGACGTGGGCCTGTCGGATGAAGAGAACTGTTTCAGGTGCACACGTTTGAAATTTTCCAGAA
AGAGCTGAATGAAAGTGAAAATTCCGTTTTCCAAGCTGTCTACGGACTGCAGAGAGCCCTGCA
GGGGGATTACAAAGATGTCGTGAACATGAAGGAGAGCAGCCGGCAGCGCCTGGAGGCCCTGAG
AGAGGCTGCAATAAAGGAAGAAACAGAAATATATGGAACCTTCTGGCAGCAGAAAAACATCAAGT
TGAAGCCCTTAAAAATATGCAACATCAAACCAAAGTTTATCCATGCTTGACGAGATTCTTGA
AGATGTAAGAAAGGCAGCGGATCGTCTGGAGGAAGAGATAGAGGAACATGCTTTTGACGACAA
TAAATCAGTCAAGGGGGTCAATTTTGAGGCAGTTCTGAGGGTGGAGGAAGAAGAGGCCAATTC
TAAGCAAAATATAACAAAACGAGAAGTGGAGGATGACTTGGGTCTTAGCATGCTGATTGACTC
CCAGAACAACAGTATATTTTGACCAAGCCCAGAGATTCAACCATCCCACGTGCAGATCACCA
CTTTATAAAGGACATTGTTACCATAGGAATGCTGTCTTGCCTTGTGGCTGGCTATGTACAGC
CATAGGATTGCCTACAATGTTTGTTTATATTATTTGTGGTGTACTTCTGGGACCTTCAGGACT
AAATAGTATTAAGTCTATTGTGCAAGTGGAGACATTAGGAGAATTTGGGGTGTCTTTTACTCT
TTTTCTTGTTGGCTTAGAATTTTCTCCAGAAAAGCTAAGAAAGGTGTGGAAGATTTCTTACA
AGGGCCGTGTTACATGACACTGTTAATGATTGCATTTGGCTTGCTGTGGGGCATCTCTTGCG
GATCAAACCCACGCAGAGCGTCTTCATTTCCACGTGTCTGTCTTGTCAAGCACACCCCTCGT
GTCCAGGTTCTCATGGGCAGTGTCTGGGGTGACAAAGAAGGCGACATTGACTACAGCACCGT
GCTCCTCGGCATGCTGGTGACGCAGGACGTGCAGCTCGGGCTCTTCATGGCCGTCATGCCGAC
TCTCATACAGGCGGGCGCCAGTGCATCTTCTAGCATTTGTCTGGAAGTTCTCCGAATCCTGGT
TTTGATTGGTCAGATTCTTTTTTCACTAGCGCGGTTTTTCTTTTATGTCTTGTTATAAAGAA
GTATCTCATTGGACCCTATTATCGGAAGCTGCACATGGAAAGCAAGGGGAACAAAGAAATCCT
GATCTTGGAATATCTGCCTTTATCTTCTTAATGTTAACGGTCACGGAGCTGCTGGACGTCTC
CATGGAGCTGGGCTGTTTCTGGCTGGAGCGCTCGTCTCCTCTCAGGGCCCCGTGGTCACCGA
GGAGATCGCCACCTCCATCGAACCCATCCGCGACTTCCTGGCCATCGTTTTCTTCGCCTCCAT
AGGGCTCCACGTGTTCCCCACGTTTGTGGCGTACGAGCTCACGGTGCTGGTGTTCCTCACCTT
GTCAGTGGTGGTGATGAAGTTTCTCCTGGCGGCGCTGGTCCTGTCTCTCATTCTGCCGAGGAG
CAGCCAGTACATCAAGTGGATCGTCTCTGCGGGGCTTGCCCAGGTGAGCGAGTTTTCTTGT
CCTGGGGAGCCGGGCGCGAAGAGCGGGCGTCTCTCGGGAGGTGTACCTCCTTATACTGAG
TGTGACCACGCTCAGCCTCTTGCTCGCCCCGGTGTGTGGAGAGCTGCAATCACGAGGTGTGT
GCCCAGACCGGAGAGACGGTCCAGCCTC**TGA**TGGCTCGGAGATGATGGACCGTGGAAGGGAAG
CGTCTGTGGGGAGTGAGCGCTTAGATGGCCAGCAGCTGCTCCTTCTGGGAAGCTCGCACCTTG
GCAACAGAACAGCCCTCTAGCAGAGCGTCAGTGCAGTCGTGTTATCCCGGCTTTTACAGAATA
TTCTTGTCTATTTTAGAATTTTCCGGAGTAGTTTATTTGCAGTCTGTTGATTATGTGCAGTA
GACCCGGGACACTGCGTTTTACCGATCACCTTGAATGTGGTGCCTGGATGTGCCTTTTTTTTT
TTTCCCTGAAATTATTATTAATTTTCTATTGTGAGTTCATCAGTTCATAGTTTTTTTAGTAAA
GAAGCAAAATTAAGGCTTTTAAAAATGTACAACTTCAGAATTATAATCTGTTAGTCAAATA
TTTGTTATTAAACATTTCTGTAATATGAAGTTGTAATCCTGGCCGTGAGCTTGGAAAGCTTACT
TTTGATTCTTAAAGCCTATGTTTTCTAAATGAGACAAATACGGATGTCTATTTGCCTTTTAT
TGTAACCTTTAAATGAAATAATTTTCATGTCAATTTCTATTAGATATATCACTTAAATATTTG
GTTTTAAATCACAAGAATATGTATTCTTTAATAAAGATAATTTATGATCATGGTAAAAAAAAA

230/550

FIGURE 230

MKVLGRSFFWVLPVLPWAVQAVEHEEVAQRVIKLHRGRGVAAMQSRQWVRDSCRKLSGLLRQ
KNAVLNKLKTAIGAVEKDVGLSDEEKLQVHTFEIFQKELNESENSVFQAVYGLQRALQGDYK
DVVNMKESSRQRLEALREAAIKEETEMYELLAAEKHQVEALKNMQHQNQSLSMLDEILEDVRK
AADRLEEEIEEHAFDDNKS VKGVNFEAVLRVEEEEANSKQNITKREVEDDLGLSMLIDSQNNQ
YILTKPRDSTIPRADHHFIKDIVTIGMLS LPCGWLCTAIGLPTMFGYIICGVLLGPSGLNSIK
SIVQVETLGEFGVFFTLFLVGLFESPEKLRKVWKISLQGPCYMTLLMIAFGLLWGHLRLRIKPT
QSVFISTCLSLSSSTPLVSRFLMGSARGDKEGDIDYSTVLLGMLVTQDVQLGLFMVMPPTLIQA
GASASSIVVEVLRILVLIGQILFSLAAVFLCLVIKKYLIGPYRKLHMESSKGNKEILILGI
SAFIFLMLTVTELLDVSMELGCFLAGALVSSQGPVVTEEIATSIEPIRDFLAIVFFASIGLHV
FPTFVAYELTVLVFLTSLVVVMKFLAALVLSLILPRSSQYIKWIVSAGLAQVSEFSFVLGSR
ARRAGVISREVYLLILSVTTLSLLLAPVLWRAAITRCVPRPERRSSL

Signal peptide:

amino acids 1-22

Transmembrane domains:amino acids 282-304, 322-337, 354-370, 379-395, 445-474, 501-520,
576-598, 641-660**N-glycosylation sites.**

amino acids 104-108, 174-178, 206-210, 230-234

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 55-59, 673-677

Tyrosine kinase phosphorylation site.

amino acids 407-414

N-myristoylation sites.amino acids 116-122, 327-333, 366-372, 401-407, 419-425, 429-435,
442-448, 525-531, 530-536**Cell attachment sequence.**

amino acids 404-407

231/550

FIGURE 231

GAGAAAAACAACAGGAAGCAGCTTACAACTCGGTGAACAACTGAGGGAACCAAACCAGAGAC
GCGCTGAACAGAGAGAATCAGGCTCAAAGCAAGTGGAAGTGGGCAGAGATTCCACCAGGACTG
GTGCAAGGCGCAGAGCCAGCCAGATTTGAGAAGAAGGCCAAAAGATGCTGGGGAGCAGAGCTG
TAATGCTGCTGTTGCTGCTGCCCTGGACAGCTCAGGGCAGAGCTGTGCCTGGGGGCAGCAGCC
CTGCCTGGACTCAGTGCCAGCAGCTTTCACAGAAGCTCTGCACACTGGCCTGGAGTGACATC
CACTAGTGGGACACATGGATCTAAGAGAAGAGGGAGATGAAGAGACTACAAATGATGTTCCCC
ATATCCAGTGTGGAGATGGCTGTGACCCCCAAGGACTCAGGGACAACAGTCAGTTCTGCTTGC
AAAGGATCCACCAGGGTCTGATTTTTTATGAGAAGCTGCTAGGATCGGATATTTTCACAGGGG
AGCCTTCTCTGCTCCCTGATAGCCCTGTGGGCCAGCTTCATGCCTCCCTACTGGGCCTCAGCC
AACTCCTGCAGCCTGAGGGTCACCACTGGGAGACTCAGCAGATTCCAAGCCTCAGTCCCAGCC
AGCCATGGCAGCGTCTCCTTCTCCGCTTCAAAATCCTTCGCAGCCTCCAGGCCTTTGTGGCTG
TAGCCGCCCCGGGTCTTTGCCCATGGAGCAGCAACCCTGAGTCCCTAAAGGCAGCAGCTCAAGG
ATGGCACTCAGATCTCCATGGCCCAGCAAGGCCAAGATAAATCTACCACCCAGGCACCTGTG
AGCCAACAGGTTAATTAGTCCATTAATTTTAGTGGGACCTGCATATGTTGAAAATTACCAATA
CTGACTGACATGTGATGCTGACCTATGATAAGGTTGAGTATTTATTAGATGGGAAGGGAAATT
TGGGGATTATTTATCCTCCTGGGGACAGTTTGGGGAGGATTATTTATTGTATTTATATTGAAT
TATGTACTTTTTTCAATAAAGTCTTATTTTTGTGGCTAAAAAAAAAAAAA

232/550

FIGURE 232

MLGSRAVMLLLLLLPWTAQGRAVPGGSSPAWTQCQQLSQKLCTLAWSAHPLVGHMDLREEGDDE
TTNDVPHIQCGDGC DPQGLRDNSQFCLQRIHQGLIFYEKL LGSDIFTGEPSLLPDSPVGQLHA
SLLGLS QLLQPEGHHWETQQIPSLSPSQPWQRLLLRFKILRSLQAFVAVAAARVFAHGAATLSP

Important features of the protein:

Signal peptide:

amino acids 1-21

Casein kinase II phosphorylation site.

amino acids 64-67

N-myristoylation sites.

amino acids 25-30, 81-86, 122-127

233/550

FIGURE 233

CCCACGCGTCCGGCCCTGTAACCAAGATACTGACTGAACATGGCTGGCGGACTCAGGCTGGGGTCTGCAGTGCAG
CATTAAATGGGCGCTGACATGAATATGGAGTAGTTTTCTCTAGCAAAGAGTAATGTGGGCCATGGAGTCAGGCCA
CCTCCTCTGGGCTCTGCTGTTTCATGCAGTCCCTGTGGCCTCAACTGACTGATGGAGCCACTCGAGTCTACTACCT
GGGCATCCGGGATGTGCAGTGGAACTATGCTCCCAAGGGAAGAAATGTCATCACGAACCAGCCTCTGGACAGTGA
CATAGTGGCTTCCAGCTTCTTAAAGTCTGACAAGAACCGGATAGGGGGAACCTACAAGAAGACCATCTATAAAGA
ATACAAGGATGACTCATACAGATGAAGTGGCCAGCCTGCCTGGTGGGCTTCCCTGGGGCCAGTGTTCAGGC
TGAAGTGGGGGATGTCATTTCTTATTCACCTGAAGAATTTTGCCACTCGTCCCTATACCATCCACCCTCATGGTGT
CTTCTACGAGAAGGACTCTGAAGGTTCCCTATACCCAGATGGCTCCTCTGGGGCACTGAAAGCTGATGACTCTGT
TCCCCGGGGGGCAGCCATATCTACAAGTGGACCATTCCAGAAGGCCATGCACCCACCGATGCTGACCAGCGTG
CCTCAGCTGGATCTACCATTTCTCATGTAGATGCTCCACGAGACATTGCAACTGGCCCTAATTGGGCCTCTCATCAC
CTGTAAAAGAGGAGCCCTGGATGGGAACCTCCCCTCCTCAACGCCAGGATGTAGACCATGATTTCTTCTCCTCTT
CAGTGTGGTAGATGAGAACCTCAGCTGGCATCTCAATGAGAACATTGCCACTTACTGCTCAGATCCGTGCTTCAGT
GGACAAGAAGATGAGACATTTTCAGGAGAGCAATAGGATGCATGCAATCAATGGCTTTGTTTTGGGAATTTACC
TGAGCTGAACATGTGTGCACAGAAACGTGTGGCCTGGCAGTTGTTTGGCATGGGCAATGAAATTGATGTCACAC
AGCATTTTTCCATGGACAGATGCTGACTACCCGTGGACACCACACTGATGTGGCTAACATCTTTCAGCCACCTT
TGTGACTGCTGAGATGGTGGCCCTGGGAACCTGGTACCTGGTTAATTAGCTGCCAAGTGAACAGTCACTTTCGAGA
TGGCATGCAGGCACTCTACAAGGTCAAGTCTTGCTCCATGGCCCTCCTGTGGACCTGCTCACAGGCAAGTTTCG
ACAGTACTTCATTGAGGCCCATGAGATTCATGGGACTATGGCCCGATGGGGCATGATGGGAGTACTGGGAAGAA
TTTGAGAGAGCCAGGCAGTATCTCAGATAAGTTTTTCCAGAAGAGCTCCAGCCGAATTGGGGGCACTTACTGGAA
AGTGGCATATGAAGCCTTTCAAGATGAGACATTTCAAGAGAAGATGCATTTGGAGGAAGATAGGCATCTTGGAA
CCTGGGGCCAGTATCCGGGCTGAGGTGGGTGACACCATTCAAGTGGTCTTCTACAACCGTGCCTCCCAGCCATT
CAGCATGCAGCCCCATGGGGTCTTTTTATGAGAAGACTATGAAGGCACTGTGTACAATGATGGCTCATCTTACCC
TGGCTTGGTTGCCAAGCCCTTTGAGAAAGTAACATACCGCTGGACAGTCCCCCTCATGCCGGTCCCAGTCTCA
GGATCCTGCTTGTCTCACTTGGATGTACTTCTCTGCTGCAGATCCCATAGAGACACAAATTCTGGCCTGGTGGG
CCCGTGTCTGGTGTGCAGGGCTGGTGCCTTGGGTGCAGATGGCAAGCAGAAAGGGGTGGATAAAGAATTCTTTCT
TCTCTTCACTGTGTGGATGAGAACAAGAGCTGGTACAGCAATGCCAATCAAGCAGCTGCTATGTTGGATTTCGG
ACTGCTTTTCAGAGGATTTGAGGGCTTCCAGACTCCAAAGCATGGATGCATGCCATTAATGGGTTTCTGTTCTTAA
CCTGCCCAGGCTGGACATGTGCAAGGGTGACACAGTGGCCTGGCACCTGCTCGGCCTGGGCACAGAGACTGATGT
GCATGGAGTCATGTTCCAGGGCAACACTGTGCAGCTTCAGGGCATGAGGAAGGGTGCAGCTATGCTCTTCTCTCA
TACCTTTGTCTATGGCCATCATGCAGCCTGACAACCTTGGGACATTTGAGATTTATGCCAGGCAGGCAGCCATCG
AGAAGCAGGGATGAGGGCAATCTATAATGTCTCCAGTGTCTTGGCCACCAAGCCACCCCTCGCCACGCTACCA
AGTGCAGAAGATCTACTATATCATGGCAGAAGAAGTAGAGTGGGACTATTGCCCTGACCCGAGCTGGGAACGGGA
ATGGCACAACCACTGTGAGAAGGACAGTTATGGTTACATTTTCTGAGCAACAAGGATGGGCTCCTGGGTTCAG
ATACAAGAAAGCTGTATTTCAGGGAATACACTGATGGTACATTCAGGATCCCTCGGCCAAGGACTGGACCAGAAGA
ACACTTGGGAATCTTGGGTCCACTTATCAAAGGTGAAGTTGGTGATATCCTGACTGTGGTATTCAAGAATAATGC
CAGCCGCCCTACTTGTGCTGCTCATGGAGTGCCTAGTAATCTACTACTGTCTGGCCACTGGCTGTGAGCCTGG
TGAGGTGGTCACTTATCATGTTGGAACATCCAGAGGACTTGGCCCTGGGCCAATGACTCTGCTTGTGTTTCTG
GATCTATTATTCTGCAGTGGATCCCATCAAGGACATGTATAGTGGCCTGGTGGGGCCCTTGGCTATCTGCCAAAA
GGGCATCCTGGAGCCCCATGGAGGACGGAGTGACATGGATCGGGAATTTGCATGTTGTTCTTGTATTTTGTATGA
AAATAAGTCTTGGTATTTGGAGGAAAATGTGGCAACCCATGGGTCCCAGGATCCAGGCAGTATTAACCTACAGGA
TGAAACTTTCTTGGAGAGCAATAAAATGCATGCAATCAATGGGAACTCTATGCCAACCTTAGGGGTCTTACCAT
GTACCAAGGAGAACGAGTGGCCTGGTACATGCTGGCCATGGGCCAAGATGTGGATCTACACACCATCCACTTTCA
TGCAGAGAGCTTCTCTATCGGAATGGCGAGAATACCGGCAGATGTGGTGGATCTGTTCCAGGGACTTTTGA
GGTGTGGAGATGGTGGCCAGCAACCCCTGGGACATGGCTGATGCACTGCCATGTGACTGACCATGTCCATGCTGG
CATGGAGACCCCTCTCACTGTTTTTCTCGAACAGAACCTTAAGCCCTCTCACCGTCATACCAAAGAGACTGA
AAAAGTGGCCCCCAGAGACATTGAAGAAGGCAATGTGAAGATGCTGGGCATGCAGATCCCCATAAAGAATGTTGA
GATGCTGGCCTCTGTTTTGTTGGCTTACCTTCTGCTCGTTGTTCTGGCTCTTGGTGGAGTGGTTTTG
GTACCAACATCGACAGAGAAAGCTACGACGCAATAGGAGGTCCATCCTGGATGACAGCTTCAAGCTTCTGTCTTT
CAAACAGTAAACATCTGGAGCCTGGAGATATCCTCAGGAAGCACATCTGTAGTGCCTCCAGCAGGCCATGGACT
AGTCACTAACCCACACTCAAAGGGGCATGGGTGGTGGAGAAGCAGAAGGAGCAATCAAGCTTATCTGGATATTT
CTTCTTTTATTTTATTTTACATGGAATAATATGATTTCACTTTTTCTTTAGTTTCTTTGCTCTACGTGGGCACCT
GGCATTAAGGGAGTACCTTATTTACATCGCAAAATTTCAACAGCTACATTTATTTCTTCTGACACTTGGGA
AGGTATTGAAATTTCTAGAAATGTATCCTTCTCACAAAGTAGAGACCAAGAGAAAACTCATTGATTGGGTTTCT
ACTTCTTTCAAGGACTCAGGAAATTTCACTTTGAAGTGGGCCAAGTGGCTGTTAAGATAACCCACACTTAAAC
TAAAGGCTAAGAATATAGGCTTGTATGGGAAATTTGAAGGTAGGCTGAGTATTGGGAATCCAAATTTGAATTTGATT
CTCCTTGGCAGTGAACACTTTGAAGAAGTGGTCAATGGGTGTTGTTGCTGCCATGAGCATGTACAACCTCTGGAGC
TAGAAGCTCCTCAGGAAAGCCAGTTCTCCAAGTTCTTAACTGTGGCACTGAAAGGAATGTTGAGTTACCTCTTC
ATGTTTTAGACAGCAACCCCTATCCATTAAAGTACTTGTAGACCAAAAAAAAAA

234/550

FIGURE 234

MWAMESGHELLWALLFMQSLWPQLTDGATRYYLIGIRDVQWNYAPKGRNVITNQPLDSDIVASS
FLKSDKNRIGGTYKKTIIYKEYKDDSYTDEVAQPAWLGLGPVLQAEVGDVILHLKNFATRPY
TIHPHGVFYEKDSEGSLYPDGSSGPLKADDSVPPGGSHIYNWTIPEGHAPTADPACLTWIIYH
SHVDAPRDIATGLIGPLITCKRGALDGNPSPQRQDVDHDFLLFSVVDENLSWHLNENIATYC
SDPASVDKEDETFQESNRMHAINGFVFGNLPFLNMCAQKRVAVHFLFGMGNEIDVHTAFFHGQM
LTTRGHHTDVANIFPATFVTAEMVPWEPGTWLI SCQVNSHFRDGMQALYKVKSCSMAPPVDLL
TGKVRQYFIEAHEIQWDYGPMGHDGSTGKNLREPGSISDKFFQKSSSRIGGTYWKVRYEAFQD
ETFOEKMHEEDRHLGILGPVIRAEVGDTIQVVFYNRASQPFMSQPHGVFYEKDYEGTVYNDG
SSYPGLVAKPFEEKVTYRWTVPFHAGPTAQDPACLTWMYFSAADPIRDTNSGLVGPLLVCRA
LGADGKQKGV DKEFFLLFTVL DENKSWYSNANQAAAMLDFRLLSE DIEGFQDSNRMHAINGFL
FSNLPRLD MCKGDTVAVHLLGLGTETDVHGVMFQGN TVQLQGM RKGAA MLFPHTFVMAIMQPD
NLGTFEIYCQAGSHREAGMRAIYNVSQC PGHQA TPQRQYQAARIYYIMAEVEWDYCPDRSWE
REWHNQSEKDSYGYIFLSNKDGLLGSR YKKAVFREYTDGTFRIPRPTGP EEHLGILGPLIKG
EVGDILT VVFKNNASRPYSVHAHGVLESTTVWPLAAEPGEVV TYQWNI PERSGPGPNDSACVS
WIIYSAVDPIKDMYSGLVG PLAICQKGILEPHGGRSDMDREFALLFLIFDENKSWYLEENVAT
HGSQDPGSINLQDET FLESNKMHAINGKLYANLRGLTMYQGERVAWYMLAMGQD VDLHTIHFH
AESFLYRNGENYRADVVDLFPGT FEVVMVASNP GTWLMHCHVT DHVHAGMETLFTVFSRTEH
LSPLTVITKETEKVPPRDIEEGNVKMLGMQIPIKNVEMLASVLVAISVTLLLVVLALGGVVWY
QHRQRKLRRNRRSILDDSFKLLSFKQ

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 1109-1130

N-glycosylation sites.amino acids 167-171, 239-243, 591-595, 717-721, 761-765, 832-836,
876-880, 934-938**Glycosaminoglycan attachment site.**

amino acids 871-875

Tyrosine kinase phosphorylation sites.

amino acids 82-90, 137-145, 494-502, 513-521

N-myristoylation sites.amino acids 212-218, 313-319, 498-504, 566-572, 672-678, 778-784,
843-849**Multicopper oxidases signature 1.**

amino acids 344-365, 696-717, 1043-1064

Multicopper oxidases signature 2.

amino acids 1048-1060

235/550

FIGURE 235

GGAAAGAGTGCTGGTACTACAACCAGGAAGTGACAGATAATGTGCTTTAAACTACATTAGAAAAGCTTCTCATAG
CAAAACTGAGAGATTGAAGCAGTGATTATTTTTACATAGTTGTCTTAAATATTTGGAGCTCTGCTGTGCATAGA
GATGGCAACATACTTAGAATACACAGCTTTCTGGGCCAGAAATTGATCTTCTGACTTTTGAGCCTTATCTGATTA
CTGCTTGGTTTCATCTTTATTTTGTAACTACTCTGTAGGCTGAAAGGGAGAGACTCTCCTTGGTTTGCAGAGCC
TGACTAGACAGGAATTCTGGCAACTGCTCCAGCAGAACTATGGCACTGAGCTAGGTTTAAATGCTGAGGAGATGG
AAAACCTTGCTACTGTGATTGAGGATGTGCAGCCAAGAAGTCCAGGAAGAAGCAGCTTGGATGACTCTGGGGAGA
GAGATGAAAAATTATCCAAGTCAATCAGTTTTACCAGTGAATCAATTAGTCGGGTTTCAGAAACAGAGTCATTTCG
ATGGAATTCATCAAAAGGAGGATTAGGCCAAAGAGGAGTCCCAAAATGAGAAACAGACCAAAAAGAGTCTCTTAC
CAACTTTGGAAAAGAAGTTAACTAGAGTGCCATCAAAGTCACTGGACTTGAATAAAAATGAATATCTTTCTCTGG
ACAAAAGCAGCACTTCAGATTCTGTTGATGAAGAAAATGTTCTGAGAAAGATCTTCATGGAAGACTTTTTATCA
ACCGTATTTTTCATATCAGTGCTGACAGAATGTTTGAATTGCTCTTACCAGTTCACGCTTTATGCAGAAATTTG
CCAGTTCTAGAAATATAATAGATGTAGTATCTACCCCTTGGACTGCAGAACTTGGAGGTGATCAGCTGAGAACGA
TGACCTACACTATAGTCTTAATAGTCCACTTACTGGAAAATGCCTGCTGCCACTGAAAAGCAGACACTGTATA
AAGAAAGTCGGGAAGCAGATTTTATTTGGTAGATTGAGAAGTACTGACACATGATGCCCCCTACCATGATTACT
TCTATACCGTGAACAGATACTGTATCATCCGATCTTCAAACAGAAATGCAGGCTAAGAGTTTCCACAGATTGGA
AATACAGAAAACAGCCATGGGGCCTTGTCAAATCTTTAATTGAAAAGAATTCTTGGAGTTCTTTGGAGGACTATT
TCAAACAGCTTGAATCAGATTGTGAATTGAAGAATCTGTATTAATCAGGCCATTGAAGACCCTGGAAAACCTTA
CTGGCCTACGAAGGAGAAGGCGAACCTTCAACCGAACAGCAGAAACAGTTCTTAACTTTCTCTCAGCATTCCT
CTGGAGATGTGGGCTTAGGTGCCAAAGGGGATATTACAGGAAAGAAAAGGAAATGGAAAACATAACGTCCTC
TTATTGTGGTAATGAGTATTTTGTGTTGTTATTAGTTTTGTTGAATGTGACACTGTTTCTGAAGCTGTCAAAGA
TAGAACATGCTGCTCAGTCTCTTTACCGTCTCCGCTCCAAGAAGAGAAATCTTTAAATTTAGCCTCTGATATGG
TGTCAAGAGCAGAAACTATTCAGAAGAATAAAGATCAGGCCATCGTTTAAAGGGAGTGCTCCGAGACTCCATAG
TGATGCTTGAACAGCTGAAGAGCTCACTCATATTGCTTCAGAAAACGTTTGATCTACTAAATAAGAATAAGACTG
GCATGGCTGTTGAAAGCTAGTGATCTGAAGGACTAAAACCGCAGAGATACTTGGAACTTAAAGAAAATACCTGGA
AGAAAACCAGACGAATGAAGGATTTTGGCATAGAACATTTCTATGTTTTTTCATTATTGAGATTTCTAATATGAA
CATTTCTTTCAGTAACATTTATTTGATAATTAGTTTCTGCTGGCCTTAATAATCCATCCTTTCACTTCTTATAGA
TATTTTTAAGCTGTGAATTTCTTCAGTGAACCATGAAATATATTATAGAATGAATTTCTCTGATACAAAAGAA
AATGACACACCCTGAATTGAGTGGTATGGTCTCATTTCTACAGTGAAGTCTGATGCTTTGTTAGCACAGAATCCG
TACATGTCCAATAGGTCGCTTTTGTAACTGAGATAAGACCAAGAGGATAAACAGGACAATATAAGAAGAAACCTC
TATGTCATTACTGATTTTTAAAGGTTCTGTTTTCAAGCATATAACATTTCCAGGTTTGTGTACTGTAAAGATTATA
ATGCTCTCATTTATTTAGCATGCAATTTAATAGTCAAACCTTTTGAATCTGCATGTTGATGATGATTATCAGAA
AGGGTCTTCTGCCATGCTGTATCTTTATGAAAGAAATAGTTGTTTTTCTTAAAGGTAAGTATCAGAGTGGGATT
ATCTTGCCCTCCTCACTTAGAATACCAACAGTCAAAGGAAGAACCATCCTCTGAGTTTAAAAACAGAAAGGTTA
TGTTAAATCTGGGCATTTAGTGACAGATCAATGCATACTTGAACCTAAGATTGGCTTCAGCTTAGCAGTCTTTC
ATGGTGGAAGTGACACATCTGGTTGAAAATAATTTGTGTATTTTTCAGTAACCATGTATGGCTTCCTTCTTTATGT
ATGTGTGTGACTTGTTTTAATTGGTAAGTTATAAGCCAGACATAGATTTTAGCTCTTTAATAAAAACCTTCAGGGG
CACGTATGTCCCAGTACAAGTGTACTGACTATCAAGTTTTAACTCAGATGCAAGCTTTGGCTCTTTCAAAAAAG
TTTTTATGCATATGTCTCCATACAAGTGGCTCATTAAATAAGAACTTTGTAACTGACTTAAATCAGATAT
TTTTTCAAGAGTTAGGGAAGTTGAAGTGTCTTACTGTTTGTCTCTTGGGCCCTTCTCTGGGGAAAAAATACA
TATCCATCTATCTATCTATATATAAACTGTGTATACATTCTTACTGTTTGAACAATATTGCCTTTAATTAATG
TTTCATTTTTCTCCAGAGTCCCCAAAGCCACATGGCATTATTATAGTCATTTTGAGATGCCTGTAGAGAATGAA
AGTATTGACTCCGTTAGAGGGAATGGGTTCTCTGGGTGAATCCAACGAAGCATACCTAGGGGTAAACGTGA
ACCTACCTGGGTTTGTGTTTTGTTTGGTAAGGATTTATGTAGTGTCTGGCTGTAAGCAAGAATGAGTGGATTATAA
ACTTGAAGATTTCTCTGTTAAAGTCACAAAATGATCGACAACAATATTTTTGTGATGTTTATTTAAACGTTGT
ATTTTATAACATACTTCAAGGAAGAGTATCGAAGTAAGTTGCTTTATAAATTAAGACTAAATTCGTATGGATGCA
GAATTCATTAATAAAATTTGAGCCTGTTACGTAAATTGAATATTAATAAAATTGAAAATTTCAAAA

236/550

FIGURE 236

MENLSLSIEDVQPRSPGRSSLDDSGERDEKLSKSI SFTSE SISRVS ETE SFDGNSSKGGLGKE
ESQNEKQTKKSLLP TLEKKL TRVPSKSLDLNKNE YLSLDKSSTSDSVDEENVPEKDLHGRLFI
NRI FHISADRMFELLFTSSRFMQKFASSRNIIDVVSTPWTAE LGGDQLRTMTYTIVLNSPLTG
KCTAATEKQ TLYKESREARFYLV DSEVLTHDVPYHDYFYTVNRYCIIRSSKQKCRLRVSTDLK
YRKQPWGLVKSLIEKNSWSSLEDYFKQLES D L L I EESVLNQAIEDPGKLTGLRRRRRTFNRTA
ETV PKLSSQHSSGDVGLGAKGDITGKKKEMENYNVT LIVVMSIFVLLLVL LNVT LFLKLSKIE
HAAQSFYRLRLQEEKSLN LASDMVSRAETIQKNKDQAHRLKGVL RDSIVMLEQLKSS LIMLQK
TFDLLNK NKTGM AVES

Transmembrane domain:

amino acids 352-371

N-glycosylation sites.

amino acids 3-7, 54-58, 312-316, 349-353, 367-371, 449-453

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 81-85, 307-311

Tyrosine kinase phosphorylation sites.

amino acids 202-211, 246-254, 341-349

N-myristoylation site.

amino acids 259-265

Amidation site.

amino acids 339-343

237/550

FIGURE 237

CAGGGGCTGGAGGGCAGGGGAGGGGATGATGTCATTCTGCTCGGCGCAATCCTGACCCTGCT
CTGGGCGCCACGGCTCAGGCTGAGGTTCTGCTGCAGCCTGACTTCAATGCTGAAAAGTTCTC
AGGCCTCTGGTACGTGGTCTCCATGGCATCTGACTGCAGGGTCTTCCTGGGCAAGAAGGACCA
CCTGTCCATGTCCACCAGGGCCATCAGGCCCACAGAGGAGGGCGGCCTCCACGTCCACATGGA
GTTCCCGGGGGCGGACGGCTGTAACCAGGTGGATGCCGAGTACCTGAAGGTGGGCTCCGAGGG
ACACTTCAGAGTCCCGGCCTTGGGCTACCTGGACGTGCGCATCGTGGACACAGACTACAGCTC
CTTCGCCGTCCTTTACATCTACAAGGAGCTGGAGGGGGCCCTCAGCACCATGGTGCAGCTCTA
CAGCCGGACCCAGGATGTGAGTCCCCAGGCTCTGAAGTCCTTCCAGGACTTCTACCCGACCCT
GGGGCTCCCCAAGGACATGATGGTCATGCTGCCCCAGTCAGATGCATGCAACCCTGAGAGCAA
GGAGGCGCCCTTGACACCTCCGGAGCCCCACCCCGCCCTTCCCAGGTGGAGCCAAAGCAGCAG
GCGCCTTTGCCCTGGAGTCAAGACCCACAGCCCTCGGGGACCACCTGGAGTCTCTCCATCCT
CCACCCCCCGCCTGTGGGATGCCTTGTGGGACGTCTCTTTCTATTCAATAAACAGATGCTGCA
GCCTCA

238/550

FIGURE 238

MMSFLLGAILTLLWAPTAQAEVLLQPDFNAEKFSGLWYVVSMASDCRVFLGKKDHLSMSTRAI
RPTEEGGLHVHMEFFPGADGCNQVDAEYLKVGSEGHFRVPALGYLDVRIVDTDYSSFAVLYIYK
ELEGALSTMVQLYSRTQDVSPQALKSFQDFYPTLGLPKDMMVMLPQSDACNPESKEAP

Signal peptide:

amino acids 1-20

Tyrosine kinase phosphorylation site.

amino acids 110-117

N-myristoylation sites.

amino acids 7-13, 79-85, 130-136

Amidation site.

amino acids 50-54

239/550

FIGURE 239

GGCGCGCTGGTCCAGGTGAGCGGGCGCGTCCCCGCGACGGCGCTGCCTGCCCCGAGGCGGTTCA
CGTAAAGACAGCGAGATCCTGAGGGCCAGCCGGAAGGAGGCGTGGATATGGAGCTGGCTGCT
GCCAAGTCCGGGGCCCCGCGCCGCTGCCTAGCGCGTCCTGGGGACTCTGTGGGGACGCGCCCCG
CGCCGCGGCTCGGGGACCCGTAGAGCCCGGCGCTGCGCGCATGGCCCTGCTCTCGCGCCCCG
GCTCACCTCCTGCTCCTCCTCATGGCCGCTGTTGTGTCAGGTGCCAGGAGCAGGCCAGACCAC
CGACTGGAGAGCCACCCTGAAGACCATCCGGAACGGCGTTTATAAGATAGACACGTACCTGAA
CGCCGCTTGGACCTCCTGGGAGGCGAGGACGGTCTCTGCCAGTATAAATGCAGTGACGGATC
TAAGCCTTTCCACGTTATGGTTATAAACCTCCCCACCGAATGGATGTGGCTCTCCACTGTT
TGGTGTTCATCTTAACATTGGTATCCCTTCCCTGACAAAGTGTGCAACCAACACGACAGGTG
CTATGAGACCTGTGGCAAAGCAAGAATGACTGTGATGAAGAATTCAGTATTGCCTCTCCAA
GATCTGCCGAGATGTACAGAAAACACTAGGACTAACTCAGCATGTTGAGGCATGTGAAACAAC
AGTGGAGCTCTTGTTTGACAGTGTTATACATTTAGGTTGTAAACCATATCTGGACAGCCAACG
AGCCGCATGCAGGTGTCATTATGAAGAAAAAAGTATCTTTTAAAGGAGATGCCGACAGCTAGT
GACAGATGAAGATGGAAGAACATAACCTTTGACAAATAACTAATGTTTTTACAACATAAACT
GTCTTATTTTTTGTGAAAGGATTATTTTGAGACCTTAAAATAATTTATATCTTGATGTTAAAC
CTCAAAGCAAAAAAGTGAGGGAGATAGTGAGGGGAGGGCACGCTTGTCTTCTCAGGTATCTT
CCCCAGCATTGCTCCCTTACTTAGTATGCCAAATGTCTTGACCAATATCAAAAACAAGTGCTT
GTTTAGCGGAGAATTTTGAAAAGAGGAATATATAACTCAATTTTCACAACCACATTTACCAAA
AAAAGAGATCAAAATATAAAATTCATCATAATGTCTGTTCAACATTATCTTATTTGGAAAATGG
GGAAATTATCACTTACAAGTATTTGTTTACTATGAAATTTTAAATACACATTTATGCCTAGAA
GGAACGGACTTTTTTTTTTCTATTTTAATTACACATAATATGTAATTAAAGTACAACATAATAT
GTTGTTTCTCTGTAGCCCGTTGAGCATATGAGTAAGTCACATTTCTATTAGGACTACTTACAA
GGACAAGGTTTCCATTTTTTCCAGTTGTAAAATTGGAACCATCAGCTGATAACCTCGTAGGGAG
CAACCCAGGATAGCTAAGTGTTATGTAATATGCCTAGAAGGTGATGTGAATGCGATTTCAGAA
GCATAGCCACTCCCATTTTATGAGCTACTCACATGACAAATGTCATCTTTTGCTATAACCTTT
GCCAAGTTAGAGAAAAGATGGATTTAATGAGATAAATGAAAAGATATTTAACCTAAAAAAA
AAAAAAAAAAAAAAAAAAAA

240/550

FIGURE 240

MALLSRPALTLLLLLMAAVVRCQEQAQTDDWRATLKTIRNGVHKIDTYLNAALDLLGGEDGLC
QYKCSDGSKPFPRYGYKPSPPNGCGSPLFGVHLNIGIPSLTKCCNQHDRCYETCGKSKNDCDE
EFQYCLSKICRDVQKTLGLTQHVQACETTVELLFDSVIHLGCKPYLDSQRAACRCHYEKTDL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites:

amino acids 57-63, 93-99

Phospholipase A2 histidine active site:

amino acids 106-114

Neuraxin and MAP1B proteins repeat proteins Block:

amino acids 109-137

241/550

FIGURE 241

GATTCCGAGCGCCTCCACTGCTGGTCCGTTGGCCAGATCAACTCGCCGCGTGGGCCGGCCGTT
CCCTGAGAGTCTGAGCGCTCGCCGCACCCCTTCCGAGCTTCTATTGGCCGTAGCAGACGTCC
GTCTGCCGCTATCTCCGCCCAATACGGAAGCGGCCTAGTCCTCCGGCTCCGACAGCTGGGTG
TCCAGGCCATGGGGCAGCCCTGGGCGGCTGGGAGCACGGACGGGGCGCCGCGCAGCTGCCTC
TCGTGCTCACCGCGCTGTGGGCCGCGGCCGTGGGCCTGGAGCTGGCTTACGTGCTGGTGCTCG
GTCCCGGGCCGCCGCCGTGGGACCCCTGGCCCGGGCCTTGACAGCTGGCGCTGGCCGCCTTCC
AGCTGCTCAACCTGCTGGGCAACGTGGGGCTCTTCCTGCGCTCGGATCCCAGCATCCGTGGCG
TGATGCTGGCCGGCCGCGGTCTGGGCCAGGGCTGGGCTTACTGCTACCAATGCCAAAGCCAGG
TGCCGCCACGCAGCGGACACTGCTCTGCCTGCCGCGTCTGCATCCTGCGTCGGGACCACCACT
GCCGCCTGCTGGGCCGCTGCGTGGGCTTCGGCAACTACCGGCCCTTCCTGTGCCTGCTGCTTC
ATGCCGCCGGCGTCCTGCTCCACGTCTCTGTGCTGCTGGGCCCTGCACTGTCGGCCCTGCTGC
GAGCCACACGCCCCCTCCACATGGCTGCCCTCCTCCTGCTTCCCTGGCTCATGTTGCTCACAG
GCAGAGTGTCTCTGGCACAGTTTGCCTTGGCCTTCGTGACGGACACGTGCGTGGCGGGTGCGC
TGCTGTGCGGGGCTGGGCTGCTCTCCATGGGATGCTGCTGCTGCGGGGCCAGACCACATGGG
AGTGGGCTCGGGGCCAGCACTCCTATGACCTGGGTCCCTGCCACAACCTGCAGGCAGCCCTGG
GGCCCCGCTGGGCCCTCGTCTGGCTCTGGCCCTTCCTGGCCTCCCCATTGCCTGGGGATGGGA
TCACCTTCCAGACCACAGCAGATGTGGGACACACAGCCTCCTGACTCCAGGAAGAGCCAGAGC
TGTGCAGGGAGGAAGGGGTGAGAGGGGGGCCCCACACCTAGACTCAGTAAGGAAGTCGGGTT
GGACCTTAACATCTGCATTGGACAACCTCACCCCTTCCTTGGCCTTGCCCCTGCCCGCCTACA
CTCCTACGTGTCCAGGGCTTGGGCCGTGACTTAGGCAGAGGAGTGCAGAGGAGGGTCTGGCAG
GGGCTGCTCAGGCCGCTAGCTGCCCCCTTTGCCAGGTTAATAAAGCACTGACTTGTTAA

242/550

FIGURE 242

MGQPWAAGSTDGAPAQLPLVLTALWAAVGLLAYVLVLGPGPPPLGPLARALQLALAAFQLL
NLLGNVGLFLRSDPSIRGVMLAGRGLGQGWAYCYQCQSQVPPRSGHCSACRVCILRRDHHCLL
LGRCVGFEGNYRPFCLLLHAAGVLLHVSVLLGPALSALLRAHTPLHMAALLLPWLMLLTGRV
SLAQFALAFVTDTCVAGALLCGAGLLFHGMLLLRGQTTWEWARGQHSYDLGPCHNLQAALGPR
WALVWLWPFLASPLPGDGITFQTTADVGHAS

Important features:**Signal peptide:**

amino acids 1-30

Transmembrane domain:

amino acids 51-66,143-160,174-191,198-214

N-myristoylation sites:

amino acids 2-8,8-14,30-36,81-87,88-94,90-96,206-212

Leucine zipper pattern:

amino acids 143-165,150-172,157-179,164-186

243/550

FIGURE 243

CTTGTCTTTGTGTCGGTTGTGATTTTCCTAATCTCTGATTTTCCTTTTCTCTCGGACGCTCTC
CCTCTTCGGACCCATTTTCTCCCGTGCTTCATGCCCTGATAGCCTGGCCCCCTCCCGGCTTCC
TTCGCTACCGGGGACGCCTCTAGTTTTTCTGAATTTCTGGCTGGCTCCACCCTCCGCGTTCAT
CTTCCTCAAGAGTTCGCCCCCTCTGGGGGCTCCTCTGTGTAATCGTCGCCTTCTCTGGGTATTT
CTGTGAACTCCGTCTCACACCATCCCGCCATCTTCTCTGCCTTGGCCCCCTTTTCTCTGTACAG
CCAGCTCTGTGTCCTTTTCTTCTCCCCCTCTAAAATCGACTCCTCTTCTCCCTGAGAGCCCCA
CCTTTGTGCCCCACTCCTCATTTTCTACGCCTCCCTCTCTCTGCTGGTCTCTCTCTCCCTG
CAAGGTTCCATTCCATCAATTTGTTTGTCTTTTGTAGGGGTGGCATCCCCTCTGACTACTGCT
CCATCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGAGGATTTCACTTCAATCTTTTCTGGT
TGCGTCTCCACTTGTA CTAGCTTGTTAGGTCCAGGTCCAGTTGTTCTGCATCTGAGGCTGGC
GTGTGCTGTCTTCTCTGATTGGCCTAATCTCCCTCACCCCCGTGAGATCTGTTGTCAGCCTTC
GTTTCTCTTTCTGTGTCCAGCTTTTCTGCGGGTCTTGGCACCTTCTTGGCCACAGATTTCT
TGGGTTACAGAGCATGTGTGTCTGAGGCATTGCAGGCAGAAAAGGGTGGCCGACGTGACCTCT
AGCTGGACTGCTGGGCAGGGGAGCTGTCTAGATAAAAATTGGAAAGAAACAGTGACCCAGAGA
CAGGTGGACAAAGAATTGCGGGACTGATGGGAACTGAGCTTGGGATCCAGACTGAAACTGATT
CCAGACTGACCTCTAGCACCCAGGACCCAGACACAGGGGCCATGGGACCCAGCATTTGAGACT
TGTGCAGCTGTTCTGCCTTCTAGGGGCCATCCCCACTCTGCCTCGGGCTGGAGCTCTTTTGTG
CTATGAAGCAACAGCCTCAAGATTCAGAGCTGTTGCTTTCCATAACTGGAAGTGGCTTCTGAT
GAGGAACATGGTGTGTAAGCTGCAAGAGGGCTGCGAGGAGACGCTAGTGTTCAATTGAGACAGG
GACTGCAAGGGGAGTTGTGGGCTTTAAAGGCTGCAGCTCGTCTTCGTCTTACCCTGCGCAAAT
CTCCTACCTTGTTTCCCCACCCGGAGTGTCATTGCCTCCTACAGTCGCGTCTGCCGGTCTTA
TCTCTGCAACAACCTCACCAATTTGGAGCCTTTTGTGAAACTCAAGGCCAGCACTCCTAAGTC
TATCACATCTGCGTCCTGTAGCTGCCCCGACCTGTGTGGGCGAGCACATGAAGGATTGCCTCCC
AAATTTTGTCAACCACTAATTCTTGCCCCCTGGGCTGCTTCTACGTGTTACAGTTCCACCTTAAA
ATTTCAAGGCAGGGTTTCTCAATACCACCTTCCCTCCTCATGGGGTGTGCTCGTGAACATAACCA
GCTTTTAGCAGATTTTCATCATATTGGGAGCATCAAAGTGACTGAGGTCCTCAACATCTTAGA
GAAGTCTCAGATTGTTGGTGCAGCATCCTCCAGGCAAGATCCTGCTTGGGGTGTGCTCTTAGG
CCTCCTGTTTGCCTTCAGGGACTTGACCATCTAGCTGCACCCGACAAGCACCCAGACTCTTTCA
CATAACAAATAAAATAGCAGAGTTCCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAA

244/550

FIGURE 244

MGPQHLRLVQLFCLLGAIPTLPRAGALLCYEATASRFRAVAFHNWKWLLMRNMVCKLQEGCEE
TLVFIETGTARGVVGFKGCSSSSSYPAQISYLVSPPGVSIASYSRVCRSYLCNNLTNLEPFVK
LKASTPKSITSASCSCPTCVGEHMKDCLPNFVTTNSCPLAASTCYSSTLKFQAGFLNTTFLLM
GCAREHNQLLADFHHIGSIKVTEVLNILEKSQIVGAASSRQDPWGVVLGLLFAFRD

Important features:**Signal peptide:**

amino acids 1-20

N-glycosylation sites:

amino acids 117-121, 183-187

N-myristoylation sites:amino acids 16-22, 25-31, 60-66, 71-77, 81-87, 100-106, 224-230,
235-241, 239-245**Prokaryotic membrane lipoprotein lipid attachment site:**

amino acids 181-192

245/550

FIGURE 245

GTGGAGTTGGGTGGTGTGCGGAGCCTCTCCCTGAGGGGCACCGCTCTTCAGGAGCTGGGCCTCCAGTGCGGCGC
GATGTCAGGCGCGGTGACAGCTCTGTGAGTCCGAGGCCGCGGCGTGGCGCTGGGCGGCTGCGGGGCTGACCGG
TCCGCTCATGGTGCCGCCACGACGCCATCGCGGGCAGGAAGGCCAGGGGTGCTGAGTTCTTCACCTCCTTTTAG
ACTGAGATCTGCCAAGTTTTCCGGCATTGCTCTTGAGGATCTCAGAAGGGCTCTTAAGACAAGACTGCAAATGGT
GTGTGTATTGTGCATGAACCGAATGAATTCCCAGAACAGTGGTTTTCACTCAGCGCAGGCGAATGGCTCTTGGGAT
TGTTATTCTTCTGCTTGTTGATGTGATATGGGTGCTTCCTCTGAACCTACTTCGTATGTTTTTACCCAGTACAA
CAAACCATTTCTTCAGCACCTTTGCAAAAACATCTATGTTTGTGTTTGTACCTTTTGGGCTTTATTATTTGGAAGCC
ATGGAGACAACAGTGTACAAGAGGACTTCGCGGAAAGCATGCTGCTTTTTTGCAGATGCTGAAGGTTACTTTGC
TGCTTGACAACAGATACAACATGAATAGTTCTTTGAGTGAACCTCTGTATGTGCCTGTGAAATTCCATGATCT
TCCAAGTAAAAACCTGAGAGCACAAACATTGATACTGAAAAAACCCCAAAAGTCTCGTGTGAGGTTCAAGTAA
TATCATGGAGATTGACAGCTTCGTCAGTCATGCAATTGGAAGCAAAGTTGTCTCGCATGTGCATATCCTGTGAA
AGAACAAGAATCCATACTGAAAACCTGTGGGAACTTACTGCAACTCAAGTAGCGAAAATTAGCTTTTTTTTTTG
CTTTGTGTGGTTTTTGGCAAATTTGTCATATCAAGAAGCACTTTCAGACACACAAGTTGCTATAGTTAATATTTT
ATCTTCAACTTCCGGACTTTTTACCTTAATCCTTGCTGCAGTATTTCCAAGTAACAGTGGAGATAGATTTACCTT
TTCTAACTATTAGCTGTAATTTTAAGCATTGGAGGCGTTGTACTGGTAAACCTGGCAGGGTCTGAAAAACCTGC
TGGAAGAGACACAGTAGGTTCCATTTGGTCTCTTGCTGGAGCCATGCTCTATGCTGTCTATATTGTTATGATTAA
GAGAAAAGTAGATAGAGAAGACAAGTTGGATATTCCAATGTTCTTTGGTTTTGTAGGTTTGTAACTGCTGCT
CTTATGGCCAGGTTTCTTTTACTTCATTATACTGGATTTGAGGACTTCGAGTTTCCAATAAAGTAGTATTAAT
GTGCATTATCATTAAATGGCCTTATTGGAACAGTACTCTCAGAGTTCCGTGTGGTTGTGGGGCTGCTTTCTTACCTC
ATCATTGATAGGCACACTTGCACTAAGCCTTACAATACCTCTGTCCATAATAGCTGACATGTGTATGCAAAAGGT
GCAGTTTTCTTGTTATTTTTTGCAGGAGCTATCCCTGTATTTTTTTCATTTTTTATGTAACTCTCCTATGCCA
TTATAATAATTGGGATCCTGTGATGGTGGGAATCAGAAGAATATTGCTTTTATATGCAGAAAACATCGAATTCA
GAGAGTTCAGAAGACAGCGAACAGTGTGAGAGTCTCATTCTATGCACAGTGTTCCTCAGGAGGATGGAGCTAG
TTAGCTGTCTGTTGTCTGTAGCCAGCTTGATAATGGAACATACAGCGAAGAGACAATCTCTGGCAAGTTTTTG
TAGAAAAATGTTTCAGTGCCTAGTCTGAAAAATAACAGTTTGAGTTCTTTGAACTCTAAATATATTTTTCTC
ATACCTGTTTTCTTCATTTTCATAATGAAGCACTTTGCTATGTAGCTGTGTACATATCACTACAGTTATAGGAAG
TTTCAGTCTACAGTCCATCCAAAGGACCAACCTGCCTTACACATCTCAAGGAATTCAGCTGTTGAAATCATTGGA
ACTAATCAAGGAATAAATCCTAATGTTCTGGGACTTTATTTTCACATGTTAAATGCTGGAATATATTATGAAAT
GTTTTCAAGAAATCACTTAAGTGTTTCATAGACCAGTATTTCTGACAGGTAAATGCTAAAATAAGCTACCTGTAA
TAAGTGTGGATTATATTTTTGGGTTTTGTAGAATATTGCAAATTAACCACACAAAAATGTTTAATTTATGCAAC
AAGCATGTTTGTGCAATTTTCATGGGACTTTAAAAAGAATAAGTATTTGAGAAAATATCTGGTTCACTTACACTA
CATTTACTGTATTATCTTTTATAGCATTAGGTGCCTTGATTTTAAATCTGTGACAAACCATGGCAAATTTTTTA
AAGGGGAAGTATTATTATAAAATGAAGAAATATGTATTTCTAAAGGCTATATTGCTGTAACTTAATTGATAAAG
CTCTGTTTAAATTTAGAGTTTTGAAGAAATAGTCTCCCTTCAATTAAGAAATTTTCATAATGGAATGATTTAAATT
GAAGTGACAAAGAGTATTATTAAAAATACAATGTTTATAAAAAA

246/550

FIGURE 246

MVPPRRHRGAGRPGVLSSSPFRLRS AKFSGIALEDLRRALKTRLQMVCFVMNRMNSQNSGF
TQRRRMALGIVILLLLVDVIWVASSELTSYVFTQYNKPFFSTFAKTSMEVLYLLGFIIWKPWRQ
QCTRGLRGKHAAFFADAEGYFAACTTDTTMNSSLSEPLYVPVKFHDLPSEKPESTNIDTEKTP
KKSRRVRFNIMEIRQLPSSHALEAKLSRMSYPVKEQESILKTVGKLTATQVAKISFFFCFVWF
LANLSYQEALSDTQVAIVNILSSTSGLFTLILAAVFPNSGDRFTLSKLLAVILSIGGVVLVN
LAGSEKPAGRDTVGSIWLAGAMLYAVYIVMIKRKVDREDKLDIPMFFGFVGLFNLLLLWPGE
FLLHYTGFEDEFEPNKVVLMCIIINGLIGTVLSEFLWLWGCFLTSSLIGTLALSLTIPLSIIA
DMCMQKVQFSWLFFAGAIPVFFSFFIVTLLCHYNNWDPVMVGIRRIFAFICRKHRIQRPEDS
EQCESLISMHSVSQEDGAS

Important features:**Transmembrane domain:**

amino acids 69-87, 105-118, 237-256, 266-285, 300-316, 332-346,
364-379, 399-419, 453-472

N-glycosylation sites:

amino acids 157-161, 255-259

N-myristoylation sites:

amino acids 14-20, 329-335, 404-410, 407-413, 418-424

247/550

FIGURE 247

CGTCTGTAGAGATATCATGAACTTCAACTTAGCTTTGGTACTTTCTTCCCTGAAGACAGAGGG
CAGAACTCTGAGTTCCAGAACCATTTTCAACTGTATTGGGGACCAATCACTTGACTCTATTCT
TGTCTCTCTGACAGATGACGCTACACTCTCCTCTGAATAATGGACACCATTTCTAAAACTGAA
TCCTGCTACTAAAATAATTAGATGATATATTTTTCCAATTCTACAATCTTGCTTTGTTTTAT
TTAGTTGTTTTCTCTCTCTCTTCCCAGTTTTCCAGAGACTGGAGCTAAACTGGGCTTTCAACA
TCATCATGAAGTTTATCCTCCTCTGGGCCCTCTTGAATCTGACTGTTGCTTTGGCCTTTAATC
CAGATTACACAGTCAGCTCCACTCCCCCTTACTTGGTCTATTTGAAATCTGACTACTTGCCCT
GCGCTGGAGTCCTGATCCACCCGCTTTGGGTGATCACAGCTGCACACTGCAATTTACCAAAGC
TTCGGGTGATATTGGGGGTACAATCCCAGCAGACTCTAATGAAAAGCATCTGCAAGTGATTG
GCTATGAGAAGATGATTCATCATCCACACTTCTCAGTCACTTCTATTGATCATGACATCATGC
TAATCAAGCTGAAAACAGAGGCTGAACTCAATGACTATGTGAAATTAGCCAACCTGCCCTACC
AAACTATCTCTGAAAATACCATGTGCTCTGTCTCTACCTGGAGCTACAATGTGTGTGATATCT
ACAAAGAGCCCGATTCACTGCAAACTGTGAACATCTCTGTAATCTCCAAGCCTCAGTGTGCGG
ATGCCTATAAAACCTACAACATCACGGAAAATATGCTGTGTGTGGGCATTGTGCCAGGAAGGA
GGCAGCCCTGCAAGGAAGTTTCTGCTGCCCCGGCAATCTGCAATGGGATGCTTCAAGGAATCC
TGTCTTTTGCGGATGGATGTGTTTTGAGAGCCGATGTTGGCATCTATGCCAAAATTTTTTACT
ATATACCCTGGATTGAAAATGTAATCCAAAATAACTTGAGCTGTGGCAGTTGTGGACCATATGA
CACAGCTTGTCCCCATCGTTCACCTTTAGAATTAAATATAAATTAACCTCCTC

248/550

FIGURE 248

MKFILLWALLNLTVALAFNPDYTVSSTPPYLVLKSDYLPAGVLIHPLWVITAHCNLPKLR
VILGVTIPADSNEKHLQVIGYEKMIHHPHFSVTSIDHDIMLIKLTAEELNDYVKLANLPYQT
ISENTMCSVSTWSYNVCDIYKEPDSLQTVNISVISKPQCRDAYKTYNITENMLCVGIVPGRRO
PCKEVSAAPAICNGMLQGILSFADGCVLRADVGIYAKIFYIIPWIENVIQNN

Important features:**Signal peptide:**

amino acids 1-17

N-glycosylation sites:

amino acids 11-15, 156-160, 173-177

Tyrosine kinase phosphorylation site:

amino acids 108-117

N-myristoylation sites:

amino acids 182-188, 203-209

Amidation site:

amino acids 185-189

Serine proteases, trypsin family, histidine active site:

amino acids 52-58

249/550

FIGURE 249

GCGAGGCGGCCGCTGTCTTCTGCTGCGGCTTCCGCGACCACAAGTACTGCTGCGACGACCCGC
ACAGCTTCTTCCCCTACGAGCACAGCTACATGTGGTGGCTCAGCATTGGCGCTCTCATAGGCC
TGTCCGTAGCAGCAGTGGTTCTTCTCGCCTTCATTGTTACCGCCTGTGTGCTCTGCTACCTGT
TCATCAGCTCTAAGCCCCACACAAAGTTGGACCTGGGCTTGAGCTTACAGACAGCAGGCCCTG
AGGAGGTTTCTCCTGACTGCCAAGGTGTGAACACAGGCATGGCGGCAGAAGTGCCAAAAGTGA
GCCCTCTCCAGCAGAGTTACTCCTGCTTGAACCCGCAGCTGGAGAGCAATGAGGGGCAGGCTG
TGAACTCCAAACGCCTCCTCCATCATTGCTTCATGGCCACAGTGACCACCAGTGACATTCCAG
GCAGCCCTGAGGAAGCCTCTGTACCCAACCCTGACCTATGTGGACCAGTCCCATAAACATTCA
ATAAATGTCTCCATACCATCAA

250/550

FIGURE 250

MWWSIGALIGLSVAADVLLAFIVTACVLCYLFISSKPHTKLDLGLSLQTAGPEEVSPDCQGV
NTGMAAEVPKVSPLQQSYSCLPQLESNEGQAVNSKRLHHCFCMATVTTSDIPGSPEEASVPN
PDL CGPVP

Important features:

Signal peptide:

Amino acids 1-26

N-myristoylation sites:

Amino acids 7-13, 11-17, 62-68, 93-99

251/550

FIGURE 251

GTGGTTTGGATTGAGCCGGGCCCCGGCCGGGGCGCCGAGTCGGAGGGGGTGGCAGTGAGCGGCG
GCAGAGGCTACGGGGCTCGGTTTGGCTGACTGGGGAGTCGGCAGGCGGCAGGAACCATGCGAG
GCCAGCGGAGCCTGCTGCTGGGCCCCGGCCCGCCTCTGCCTCCGCCTCCTTCTGCTGCTGGGT
ACAGGCGCCGCTGTCCACCTCTACTCCGGGGTCTAGTACAGCGCTGGCGCTACGGCAAGGTCT
GCCTGCGCTCCCTGCTCTACAACTCCTTTGGGGGCGAGTGACACCGCTGTTGATGCTGCCTTTG
AGCCTGTCTACTGGCTGGTAGACAACGTGATCCGCTGGTTTGGAGTGGTGTTCGTGGTCCTGG
TGATCGTGCTGACAGGCTCCATTGTAGCTATCGCCTACCTGTGTGTCCTGCCTCTCATCCTCC
GAACCTACTCAGTGCCACGACTCTGCTGGCATTCTTTCTATAGCCACTGGAATCTGATCCTGA
TTGTCTTCCACTACTACCAGGCCATCACCCTCCGCCTGGGTACCCACCCAGGGCAGGAATG
ATATCGCCACCGTCTCCATCTGTAAGAAGTGCAATTTACCCCAAGCCAGCCCGAACACACCCT
GCAGCATCTGCAACAGGTGTGTGCTGAAGATGGATCACCCTGCCCCTGGCTAAACAATTGTG
TGGGCCACTATAACCATCGGTACTTCTTCTCTTTCTGCTTTTTTCATGACTCTGGGCTGTGTCT
ACTGCAGCTATGGAAGTTGGGACCTTTTCCGGGAGGCTTATGCTGCCATTGAGACTTATCACC
AGACCCACCCACCCACCTTCTCCTTTGAGAAAGGATGACTCACAAGAGTCTTGTCTACCTCT
GGTTCCTGTGCAGTTCTGTGGCACTTGCCCTGGGTGCCCTAACTGTATGGCATGCTGTTCTCA
TCAGTCGAGGTGAGACTAGCATCGAAAGGCACATCAACAAGAAGGAGAGACGTCGGCTACAGG
CCAAGGGCAGAGTATTTAGGAATCCTTACAACCTACGGCTGCTTGGACAACCTGGAAGGTATTCC
TGGGTGTGGATACAGGAAGGCACTGGCTTACTCGGGTGCTCTTACCTTCTAGTCACTTGCCCC
ATGGGAATGGAATGAGCTGGGAGCCCCCTCCCTGGGTGACTGCTCACTCAGCCTCTGTGATGG
CAGTGTGAGCTGGACTGTGTGAGCCACGACTCGAGCACTCATTCTGCTCCCTATGTTATTTCA
AGGGCCTCCAAGGGCAGCTTTTCTCAGAATCCTTGATCAAAAAGAGCCAGTGGGCCTGCCTTA
GGGTACCATGCAGGACAATTCAAGGACCAGCCTTTTTTACCACTGCAGAAGAAAGACACAATGT
GGAGAAATCTTAGGACTGACATCCCTTTACTCAGGCAAACAGAAAGTTCCAACCCAGACTAGG
GGTCAGGCAGCTAGCTACCTACCTTGCCCGAGTGCTGACCCGGACCTCCTCCAGGATACAGCAC
TGGAGTTGGCCACCACCTCTTCTACTTGCTGTCTGAAAAACACCTGACTAGTACAGCTGAGA
TCTTGGCTTCTCAACAGGGCAAAGATACCAGGCCTGCTGCTGAGGTCACTGCCACTTCTCACA
TGCTGCTTAAGGGAGCACAAATAAAGGTATTGATTTTTTAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAA

252/550

FIGURE 252

MRGQRSLLLGPARLCLRLLLLLGYYRRRCPLLRLGLVQRWRYGKVCLRSLLYNSFGGSDTAVDA
AFEPVYWLVDNVIRWFGVVFVVLVIVLTGSIVAIAYLCVLPLILRTYSVPRLCWHFFYSHWNL
ILIVFHYYQAITTPPGYPPQGRNDIATVSICKKCIYPKPARTHHCSICNRCVLKMDHHC PWLN
NCVGHYNHRYFFSFCFFMTLGCVCYCSYGSWDLFREAYAAIETYHQTPPPTFSFRERMTHKSLV
YLWFLCSSVALALGALTVWHAVLISRGETSIERHINKKERRRLQAKGRVFRNPYNYGCLDNWK
VFLGVDTGRLHLTRVLLPSSHLPHGNGMSWEPPPWVTAHSASVMAV

Important features:**Transmembrane domain:**

amino acids 88-100, 202-216, 254-274

N-myristoylation sites:

amino acids 55-61, 56-62, 92-98, 210-216, 309-315, 319-325, 340-346

Prokaryotic membrane lipoprotein lipid attachment site:

amino acids 201-212

253/550

FIGURE 253

GATCAAGCGCCTTCCTTTCCCTTCCTCTCCCTACTTGGCCTTTGCCCTAAGCCAAGACCTGGCCATCAGCCTGGC
TGCAGGGGCTGCAGAGCCAGCTGCACTTTTTCAGGTATGGGGGAGGGCCAGGCACCATGAAGCCAGTGTGGGTC
GCCACCTTCTGTGGATGCTACTGCTGGTGGCCAGGCTGGGGGCCGCCGGAAGGGGTCCCCAGAAGAGGCCTCC
TTCTACTATGGAACCTTCCCTCTTGGCTTCTCCTGGGGCGTGGGCAGTTCTGCCTACCAGACGGAGGGCGCCTGG
GACCAGGACGGGAAAGGGCCTAGCATCTGGGACGTCTTCACACACAGTGGGAAGGGGAAAGTGCTTGGGAATGAG
ACGGCAGATGTAGCCTGTGACGGCTACTACAAGGTCCAGGAGGACATCATTCTGCTGAGGGAAGTGCACGTCAAC
CACTACCGATTCTCCCTGTCTTGGCCCCGGCTCCTGCCACAGGCATCCGAGCCGAGCAGGTGAACAAGAAGGGA
ATCGAATTCTACAGTGATCTTATCGATGCCCTTCTGAGCAGCAACATCACTCCCATCGTGACCTTGCACTACTGG
GATCTGCCACAGCTGCTCCAGGTCAAATACGGTGGGTGGCAGAATGTGAGCATGGCCAACTACTTCAGAGACTAC
GCCAACCTGTGCTTTGAGGCCTTTGGGGACCGTGTGAAGCACTGGATCACGTTCAGTGATCCTCGGGCAATGGCA
GAAAAAGGCTATGAGACGGGGCCACCATGCGCCGGGCCTGAAGCTCCGCGGCACCGGCCTGTACAAGGCAGCACAC
CACATCATTAAAGGCCACGCCAAAACCTGGCATTCTTATAACACCACGTGGCGCAGCAAGCAGCAAGGTCTGGTG
GGAATTTCACTGAAGTGTGACTGGGGGAACCTGTGGACATTAGTAACCCCAAGGACCTAGAGGCTGCCGAGAGA
TACCTACAGTTCTGTCTGGGCTGGTTTGCCAACCCCATTTATGCCGGTGACTACCCCAAGTCATGAAGGACTAC
ATTGGAAGAAAGAGTGACAGCAAGGCCTGGAGATGTCGAGGTTACCGGTGTTCTCACTCCAGGAGAAGAGCTAC
ATTAAAGGCACATCCGATTTCTTGGGATTAGGTCATTTTACTACTCGGTACATCACGGAAGGAACTACCCCTCC
CGCCAGGGGGCCAGCTACCAGAACGATCGTGACTTGATAGAGCTGGTTGACCCAACTGGCCAGATCTGGGGTCT
AAATGGCTATATTCTGTGCCATGGGGATTTAGGAGGCTCCTTAACCTTGCTCAGACTCAATACGGTGATCCTCCC
ATATATGTGATGAAAAATGGAGCATCTCAAAAATCCACTGTACTCAATTATGTGATGAGTGGAGAATTCAATAC
CTTAAAGGATACATAAATGAAATGCTAAAAGCTATAAAAGATGGTGCTAATATAAAGGGGTATACTTCCTGGTCT
CTGTTGGATAAGTTTGAATGGGAGAAAGGATACTCAGATAGATATGGATTCTACTATGTTGAATTTAACGACAGA
AATAAGCCTCGCTATCCAAAGGCTTCAGTTCAATATTACAAGAAGATTATCATTGCCAATGGGTTTCCCAATCCA
AGAGAGGTGGAAGTTGGTACCTCAAAGCTTTGGAACTTGCTCTATCAACAATCAGATGCTTGCTGCAGAGCCT
TTGCTAAGTCACATGCAAATGGTTACGGAGATCGTGGTACCCACTGTCTGCTCCCTCTGTGTCCTCATCACTGCT
GTTCTACTAATGCTCCTCCTGAGGAGGCAGAGCTTGAGACAGGATTATCAATTTTGGAGCTTCATAAGAGAATCTT
CAGGATCTTCCTCCCTTTTCTGCTTTGAGGGTTCCATACATTGCTGTTTTTCAGGTTCTACAATAATTACCTTTT
TTTCTCTTCTCTTTTTGGCTTGTGCTGGGATTTAAGAATTAGAAAATAAAAATAAGCAGAAATTA

254/550

FIGURE 254

MKPVVWATLLWMLLLVPRLLGAARKGSPEEASFYYGTFPLGFSWGVGSSAYQTEGAWDQDGKGPSIWDVFTHSKGK
KVLGNETADVACDGYKVEDIILLRELHVNHYRFSLSWPRLPTGIRAEQVNKKGIEFYSDLIDALLSSNITPI
VTLHHWDLPQLLQVKYGGWQNVSMANYFRDYANLCFEAFGDRVKHWITFSDPRAMAEKGYETGHHAPGLKLRGTG
LYKAAHHIIKAHAKTWHSYNTTWRSKQQGLVGISLNCDWGEPVDISNPKDLEAAERYLQFCLGWFFANPIYAGDYP
QVMKDYIGRKSAEQGLEMSRLPVFSLQEKSYIKGTSDFLGLGHFTTRYITERNYPSRQGPSYQNDRLIELVDPN
WPDLGSKWLYSVPWGFRRLLNFAQTQYGDPPYYVMENGASQKFHCTQLCDEWRIQYLKGYINEMLKAIKDGANIK
GYTSWSLLDKFEWEGYSRDYGFYYVEFNDNRNPKPRYPKASVQYYKKIIANGFPNPREVESWYLKALETCSINNQ
MLAAEPLLSHMOMVTEIVVPTVCSLCVLITAVLLMLLLRRQS

Important features:**Signal peptide:**

amino acids 1-21

Transmembrane domain:

amino acids 541-558

N-glycosylation sites:

amino acids 80-84, 171-175, 245-249

Glycosaminoglycan attachment site:

amino acids 72-76

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

amino acids 23-27, 564-568

Tyrosine kinase phosphorylation sites:

amino acids 203-211, 347-355, 460-468, 507-514

N-myristoylation sites:

amino acids 44-50, 79-85, 167-173, 225-231, 257-263, 315-321

Amidation site:

amino acids 307-311

Glycosyl hydrolases family 1 active site:

amino acids 407-416

Glycosyl hydrolases family 1 N-terminal signature:

amino acids 41-56

Motif name Glycosyl hydrolases family:

amino acids 37- 67

255/550

FIGURE 255

CGCGAAGATGCGAAAGGTGGTTTTGATCACCGGGGCTAGCAGTGGCATTGGCCTGGCCCTCTG
CAAGCGGCTGCTGGCGGAAGATGATGAGCTTCATCTGTGTTTGGCGTGCAGGAACATGAGCAA
GGCAGAAGCTGTCTGTGCTGCTCTGCTGGCCTCTCACCCCACTGCTGAGGTCACCATTGTCCA
GGTGGATGTCAGCAACCTGCAGTCGGTCTTCCGGGCTCCAAGGAACTTAAGCAAAGGTTTCA
GAGATTAGACTGTATATATCTAAATGCTGGGATCATGCCTAATCCACAATAAATATCAAAGC
ACTTTTCTTTGGCCTCTTTTCAAGAAAAGTGATTTCATATGTTCTCCACAGCTGAAGGCCTGCT
GACCCAGGGTGATAAGATCACTGCTGATGGACTTCAGGAGGTGTTTGAGACCAATGTCTTTGG
CCATTTTATCCTGATTTCGGAACTGGAGCCTCTCCTCTGTACAGTGACAATCCATCTCAGCT
CATCTGGACATCATCTCGCAGTGCAAGGAAATCTAATTTAGCCTCGAGGACTTCCAGCACAG
CAAAGGCAAGGAACCCTACAGCTCTTCCAAATATGCCACTGACCTTTTGAGTGTGGCTTTGAA
CAGGAACTTCAACCAGCAGGGTCTCTATTCCAATGTGGCCTGTCCAGGTACAGCATTGACCAA
TTTGACATATGGAATTCTGCCTCCGTTTATATGGACGCTGTTGATGCCGGCAATATTGCTACT
TCGCTTTTTTTGCAAATGCATTCACTTTGACACCATATAATGGAACAGAAGCTCTGGTATGGCT
TTTCCACCAAAAGCCTGAATCTCTCAATCCTCTGATCAAATATCTGAGTGCCACCACTGGCTT
TGGAAGAAATTATATTATGACCCAGAAGATGGACCTAGATGAAGACACTGCTGAAAAATTTTA
TCAAAGTTACTGGAAGTGGAAAAGCACATTAGGGTCACTATTCAAAAAACAGATAATCAGGC
CAGGCTCAGTGGCTCATGCCTATAATTCCAGCACTTTGGGAGGCCAAGGCAGAAGGATCACTT
GAGACCAGGAGTTCAAGACCAGCCTGAGAAACATAGTGAGCCCTTGTCTCTACAAAAGAAAT
AAAAATAATAGCTGGGTGTGGTGGCATGCGCATGTAGTCCCAGCTACTCAGAAGGATGAGGTG
GGAGGATCTCTTGAGGCTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGTGCCACTGCACTCC
AGCCTGGGTGACAGCGAGACCCTGTCTCAAAATATGTATATATTTAATATATATATAAAACCA
GAGCTGACAATGACACTCTGGAACATTGCATACCTTCTGTACATTCTGGGGTACATGGATTTC
TACTGAGTTGGATAATATGCATTTGTAATAAACTATGAACATATGAA

256/550

FIGURE 256

MRKVVLITGASSGIGLALCKRLLAEDDELHLCLACRNMSKAEAVCAALLASHPTAEVTIVQVD
VSNLQSVFRASKELKQRFQRLDCIYLNAGIMPNPQLNIKALFFGLFSRKVIHMFSTAEGLLTQ
GDKITADGLQEVFETNVFGHFILIRELEPLLCHSDNPSQLIWTSSRSARKSNFSLEDFQHSGK
KEPYSSSKYATDLLSVALNRNFNQQGLYSNVACPGTALTNLTYGILPPFIWTLLMPAILLLRF
FANAFTLTPYNGTEALVWLFHQKPESLNPLIKYLSATTGFGGRNYIMTQKMDLDEDTAEKFYQK
LLELEKHIRVTIQKTDNQARLSGSCL

Important features:**Transmembrane domain:**

amino acids 234-254

N-glycosylation sites:

amino acids 37-41, 178-182, 229-233, 263-267

Glycosaminoglycan attachment site:

amino acids 12-16

N-myristoylation sites:

amino acids 9-15, 13-19, 15-21, 215-221, 224-230

257/550

FIGURE 257

CGGACGCGTGGGGCCGTATGCGCGGCTCTGTGGAGTGCACCTGGGGTTGGGGGCACTGTGCCC
CCAGCCCCCTGCTCCTTTGGACTCTACTTCTGTTTGCAGCCCCATTTGGCCTGCTGGGGGAGA
AGACCCGCCAGGTGTCTCTGGAGGTCATCCCTAACTGGCTGGGCCCCCTGCAGAACCTGCTTC
ATATACGGGCAGTGGGCACCAATTCCACACTGCACTATGTGTGGAGCAGCCTGGGGCCTCTGG
CAGTGGTAATGGTGGCCACCAACACCCCCACAGCACCTGAGCATCAACTGGAGCCTCCTGC
TATCCCCTGAGCCCGATGGGGGCTGATGGTGCTCCCTAAGGACAGCATTCAGTTTTCTTCTG
CCCTTGTTTTTACCAGGCTGCTTGAGTTTGACAGCACCAACGTGTCCGATACGGCAGCAAAGC
CTTTGGGAAGACCATATCCTCCATACTCCTTGCCGATTTCTCTTGGAACAACATCACTGATT
CATTGGATCCTGCCACCCTGAGTGCCACATTTCAAGGCCACCCATGAACGACCCTACCAGGA
CTTTTGCCAATGGCAGCCTGGCCTTCAGGGTCCAGGCCTTTTCCAGGTCCAGCCGACCAGCCC
AACCCCTCGCCTCCTGCACACAGCAGACACCTGTCAGCTAGAGGTGGCCCTGATTGGAGCCT
CTCCCCGGGGAAACCGTTCCCTGTTTTGGGCTGGAGGTAGCCACATTGGGCCAGGGCCCTGACT
GCCCCTCAATGCAGGAGCAGCACTCCATCGACGATGAATATGCACCGGCCGTCTTCCAGTTGG
ACCAGCTACTGTGGGGCTCCCTCCCATCAGGCTTTGCACAGTGGCGACCAGTGGCTTACTCCC
AGAAGCCGGGGGGCCGAGAATCAGCCCTGCCCTGCCAAGCTTCCCCTCTTCATCCTGCCTTAG
CATACTCTCTTCCCCAGTCACCCATTGTCCGAGCCTTCTTTGGGTCCCAGAATAACTTCTGTG
CCTTCAATCTGACGTTTCGGGGCTTCCACAGGCCCTGGCTATTGGGACCAACACTACCTCAGCT
GGTCGATGCTCCTGGGTGTGGGCTTCCCTCCAGTGGACGGCTTGTCCCCACTAGTCCTGGGCA
TCATGGCAGTGGCCCTGGGTGCCCCAGGGCTCATGCTGCTAGGGGGCGGCTTGGTTCTGCTGC
TGACCACAAGAAGTACTCAGAGTACCAGTCCATAAATTAAGCCCCGCTCTCTGGAGGGAAGG
ACATTACTGAACCTGTCTTGCTGTGCCTCGAAACTCTGGAGGTTGGAGCATCAAGTTCCAGCC
GGCCCTTCACTCCCCCATCTTGCTTTTCTGTGGAACCTCAGAGGCCAGCCTCGACTTCCTGG
AGACCCCCAGGTGGGGCTTCCTTCATACTTTGTTGGGGGACTTTGGAGGCGGGCAGGGGACAG
GGCTATTGATAAGGTCCCCTTGGTGTTGCCTTCTTGCATCTCCACACATTTCCCTTGATGGG
ACTTGCAGGCCTAAATGAGAGGCATTCTGACTGGTTGGCTGCCCTGGAAGGCAAGAAAATAGA
TTTATTTTTTTTCACAGGGAAAAAAAAAAAAA

258/550

FIGURE 258

MRGSVECTWGWGHCAPSPLLLWTLTLLFAAPFGLLGKTRQVSLEVIPNWLGPLQNLLHIRAVG
TNSTLHYVWSSLGPLAVVMVATNTPHSTLSINWSLLLSPEPDGGLMVLPKDSIQFSSALVFTR
LLEFDSTNVSDTAAPLGRPYPPYSLADFSWNNITDSLDPATLSATFQGHMNDPTRTFANGS
LAFRVQAQFSRSSRPAQPPRLHTADTCQLEVALIGASPRGNRSLEFGLEVATLGQGPDCPSMQE
QHSIDDEYAPAVFQLDQLLWGSPLSGFAQWRPVAYSQKPGGRESALPCQASPLHPALAYSLPQ
SPIVRAFFGSQNNFCAFNLTFGASTGPGYWDQHYLSWSMLLGVGFPVPDGLSPLVLGIMAVAL
GAPGLMLLGGGLVLLLHHKKYSEYQSIN

Important features:**Signal peptide:**

amino acids 1-35

Transmembrane domain:

amino acids 365-386

N-glycosylation sites:

amino acids 65-69, 95-99, 134-138, 159-163, 187-191, 230-234, 333-337

cAMP- and cGMP-dependent protein kinase phosphorylation site:

amino acids 397-401

N-myristoylation sites:

amino acids 3-9, 63-69, 235-241, 273-279, 292-298, 324-330

Leucine zipper pattern:

amino acids 371-393

259/550

FIGURE 259

CAGGCGGGCCCCCGCGCGCAGGGCCCTGGACCCGCGCGGCTCCCGGGGATGGTGAGCAAGGCGCTGCTGCGCCT
CGTGTCTGCCGTCAACCGCAGGAGGATGAAGCTGCTGCTGGGCATCGCCTTGCTGGCCTACGTGCGCTCTGTTTG
GGGCAACTTCGTTAATATGAGGTCTATCCAGGAAAATGGTGAACATAAAATTGAAAGCAAGATTGAAGAGATGGT
TGAACCACTAAGAGAGAAAATCAGAGATTTAGAAAAAGCTTTACCCAGAAATACCCACCAGTAAAGTTTTTATC
AGAAAAGGATCGGAAAAGAATTTTGATAACAGGAGGCGCAGGGTTTCGTGGGCTCCCATCTAACTGACAACTCAT
GATGGACGGCCACGAGGTGACCGTGGTGGACAATTTCTTCACGGGCAGGAAGAGAAACGTGGAGCACTGGATCGG
ACATGAGAACTTCGAGTTGATTAAACACGACGTGGTGGAGCCCCCTCTACATCGAGTTGACCAGATATACCATCT
GGCATCTCCAGCCTCCOCTCCAACTACATGTATAATCCTATCAAGACATTAAAGACCAATACGATTGGGACATT
AAACATGTTGGGGCTGGCAAAACGAGTCGGTGCCGCTGCTCCTGGCCTCCACATCGGAGGTGTATGGAGATCC
TGAAGTCCACCTCAAAGTGAGGATTACTGGGGCCACGTGAATCCAATAGGACCTCGGGCCTGCTACGATGAAGG
CAAACGTGTTGCAGAGACCATGTGCTATGCCTACATGAAGCAGGAAGGCGTGGAAGTGCGAGTGGCCAGAATCTT
CAACACCTTTGGGCCACGCATGCACATGAACGATGGGCGAGTAGTCAGCAACTTCATCCTGCAGGCGCTCCAGGG
GGAGCCACTCACGGTATACGGATCCGGGTCTCAGACAAGGGCGTTCCAGTACGTCAGCGATCTAGTGAATGGCCT
CGTGGCTCTCATGAACAGCAACGTCAGCAGCCCGGTCAACCTGGGGAACCCAGAAGAACACACAATCCTAGAATT
TGCTCAGTTAATTA AAAACCTTGTTGGTAGCGGAAGTGAAATTCAGTTTCTCTCCGAAGCCCAGGATGACCCACA
GAAAAGAAAACCAGACATCAAAAAGCAAAGCTGATGCTGGGGTGGGAGCCCGTGGTCCCCTGGAGGAAGGTTT
AAACAAAGCAATTCACTACTTCCGTAAAGAACTCGAGTACCAGGCAAATAATCAGTACATCCCCAAACCAAAGCC
TGCCAGAATAAAGAAAGGACGGACTCGCCACAGCTGAACTCCTCACTTTTAGGACACAAGACTACCATTGTACAC
TTGATGGGATGTATTTTTGGCTTTTTTTGTTGTCGTTTAAAGAAAGACTTTAACAGGTGTCATGAAGAACAAC
TGGAATTTCACTCTGAAGCTTGCTTTAATGAAATGGATGTGCCTAAAGCTCCCCTCAAAAACTGCAGATTTTG
CCTTGCACTTTTGAATCTCTCTTTTATGTAAATAGCGTAGATGCATCTCTGCGTATTTCAAGTTTTTTTAT
CTTGCTGTGAGAGCATATGTTGTGACTGTCGTTGACAGTTTTATTTACTGGTTTCTTTGTGAAGCTGAAAAGGAA
CATTAAAGCGGGACAAAAAATGCCGATTTTATTTATAAAAGTGGGTACTTAATAAATGAGTCGTTATACTATGCAT
AAAGAAAAATCCTAGCAGTATTGTCAGGTGGTGGTGCGCCGGCATTGATTTTAGGGCAGATAAAAGAATTCTGTG
TGAGAGCTTTATGTTTCTCTTTTAATTCAGAGTTTTTCCAAGGTCTACTTTTGAGTTGCAAACTTGACTTTGAAA
TATTCCTGTTGGTCATGATCAAGGATATTTGAAATCACTACTGTGTTTGTCTGCGTATCTGGGGCGGGGCAGGT
TGGGGGGCACAAAGTTAACATATTCTTGTTAACCATGGTTAAATATGCTATTTTAATAAAATATTGAACTCA

260/550

FIGURE 260

MVSKALLRLVSAVNRRRMKLLLGIALLAYVASVWGNFVNMRSIQENGELKIESKIEEMVEPLR
EKIRDLEKSFTQKYPPVKFLSEKDRKRILITGGAGFVGSHLTDKLMDGHEVTVVDNFFTGRK
RNVEHWIGHENFELINHADVVEPLYIEVDQIYHLASPPNYMYPNPIKTLKTNTIGTLNMLGL
AKRVGARLLLASTSEVYGDPEVHPQSEDYWGHNPIGPRACYDEGKRVAETMCYAYMKQEGVE
VRVARIFNTFGPRMHMNDGRVVSNFILQALQGEPLTVYSGSQTRAFQYVSDLVNLGLVALMNS
NVSSPVNLGNPEEHTILEFAQLIKNLVSGSGSEIQFLSEAQDDPQKRKPDIKKAKLMLGWEPVV
PLEEGLNKAIHYFRKELEYQANNQYIPKPKPARIKKGRTRHS

Important features:**Signal peptide:**

amino acids 1-32

N-glycosylation site:

amino acids 316-320

Tyrosine kinase phosphorylation site:

amino acids 235-244

N-myristoylation sites:

amino acids 35-41, 101-107, 383-389

Amidation sites:

amino acids 123-127, 233-237

261/550

FIGURE 261

GCGTGGTGCGGGGGCGTGGGGAAATCGGGTTGCCCCAGCCGTTACTGGTCCGCGCAGTCAGGG
CATCCTCCGCATCCTCCACATCCTTCC**ATG**GGCTCTGAAGAATAAATTCAGTTGTTTATGGATC
TTGGGTCTGTGTTTGGTAGCCACTACATCTTCCAAAATCCCATCCATCACTGACCCACACTTT
ATAGACAACTGCATAGAAGCCCACAACGAATGGCGTGGCAAAGTCAACCCCTCCCGCGGCCGAC
ATGAAATACATGATTTGGGATAAAGGTTTAGCAAAGATGGCTAAAGCATGGGCAAACCAGTGC
AAATTTGAACATAATGACTGTTTGGATAAATCATATAAATGCTATGCAGCTTTTGAATATGTT
GGAGAAAATATCTGGTTAGGTGGAATAAAGTCATTACACCAAGACATGCCATTACGGCTTGG
TATAATGAAACCAATTTTATGATTTTGATAGTCTATCATGCTCCAGAGTCTGTGGCCATTAT
ACACAGTTAGTTTGGGCCAATTCATTTTATGTCGGTTGTGCAGTTGCAATGTGTCCTAACCTT
GGGGGAGCTTCAACTGCAATATTTGTATGCAACTACGGACCTGCAGGAAATTTTGCAATATG
CCTCCTTACGCAAGAGGAGAATCTTGCTCTCTCTGCTCAAAGAAGAGAAATGTGTAAAGAAC
CTCTGCAGGACTCCACAACCTTATTATACCTAACCAAAATCCATTTCTGAAGCCAACGGGGAGA
GCACCTCAGCAGACAGCCTTTAATCCATTCAGCTTAGGTTTTCTTCTTCTGAGAATCTTT**TAA**
TGTCAATTTATATACAAAAGAAATTCTCAAATGTTAAAATAAAGGAATAGTTTATTGCTTAATA

262/550

FIGURE 262

MALKNKFSLWILGLCLVATTSSKIPSITDPHFIDNCIEAHNEWRGKVNPPAADMKYMIWDKG
LAKMAKAWANQCKFEHNDCLDKSYKCYAAFEYVGENIWLGGIKSFTPRHAITAWYNETQFYDF
DSLSCSRVCGHYTQLVWANSFYVGCAVAMCPNLGGASTAIFVCNYGPAGNFANMPPYARGESC
SLCSKEEKCCKVKNLCRTPQLIIPNQNPFLKPTGRAPQQTAFNPFSLGFLLLRIF

Important features:**Signal peptide:**

amino acids 1-23

N-glycosylation site:

amino acids 119-123

N-myristoylation sites:

amino acids 103-109,150-156,160-166,161-167,175-181

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1:

amino acids 136-156

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2:

amino acids 166-178

263/550

FIGURE 263

CGCCCTCCGACCCGCCCCGCGGCGCATTGTGGGATCTGTCCGGCTTGTCAAGGTGGTGGAGGAAA
AGGCGCTCCGTCATGGGGATCCAGACGAGCCCCGTCTGCTGGCCTCCCTGGGGGTGGGGCTG
GTCACCTCTGCTCGGCCTGGCTGTGGGCTCCTACTTGGTTCGGAGGTCCCGCCGGCCTCAGGTC
ACTCTCCTGGACCCCAATGAAAAGTACCTGCTACGACTGCTAGACAAGACGACTGTGAGCCAC
AACACCAAGAGGTTCCGCTTTGCCCTGCCCACCGCCCACCACACTCTGGGGCTGCCTGTGGGC
AAACATATCTACCTCTCCACCCGAATTGATGGCAGCCTGGTCATCAGGCCATACACTCCTGTC
ACCAGTGATGAGGATCAAGGCTATGTGGATCTTGTCAATCAAGGTCTACCTGAAGGGTGTGCAC
CCCAAATTTCTGAGGGAGGGAAGATGTCTCAGTACCTGGATAGCCTGAAGGTTGGGGATGTG
GTGGAGTTTCGGGGGCCAAGCGGGTTGCTCACTTACACTGGAAAAGGGCATTTTAACATTTCAG
CCCAACAAGAAATCTCCACCAGAACCCCGAGTGGCGAAGAACTGGGAATGATTGCCGGCGGG
ACAGGAATCACCCCAATGCTACAGCTGATCCGGGCCATCCTGAAAGTCCCTGAAGATCCAACC
CAGTGCTTTCTGCTTTTTGCCAACCAGACAGAAAAGGATATCATCTTGCGGGAGGACTTAGAG
GAACTGCAGGCCCGCTATCCCAATCGCTTTAAGCTCTGGTTCACCTCTGGATCATCCCCCAAAA
GATTGGGCCTACAGCAAGGGCTTTGTGACTGCCGACATGATCCGGGAACACCTGCCCCGCTCCA
GGGGATGATGTGCTGGTACTGCTTTGTGGGCCACCCCAATGGTGCAGCTGGCCTGCCATCCC
AACTTGACAAACTGGGCTACTCACAAAAGATGCGATTACCTACTTGAGCATCCTCCAGCTTC
CCTGGTGCTGTTTCGCTGCAGTTGTTCCCATCAGTACTCAAGCACTATAAGCCTTAGATTCCCT
TTCCTCAGAGTTTCAGGTTTTTTTCAGTTACATCTAGAGCTGAAATCTGGATAGTACCTGCAGG
AACAATATTCTGTAGCCATGGAAGAGGGCAAGGCTCAGTCACTCCTTGGATGGCCTCCTAAA
TCTCCCCGTGGCAACAGGTCCAGGAGAGGCCCATGGAGCAGTCTCTTCCATGGAGTAAGAAGG
AAGGGAGCATGTACGCTTGGTCCAAGATTGGCTAGTTCCCTTGATAGCATCTTACTCTCACCTT
CTTTGTGTCTGTGATGAAAGGAACAGTCTGTGCAATGGGTTTTACTTAACTTCACTGTTCAA
CCTATGAGCAAATCTGTATGTGTGAGTATAAGTTGAGCATAGCATACTTCCAGAGGTGGTNTT
ATGGAGATGGCAAGAAAGGAGGAAATGATTTCTTCAGATNTCAAAGGAGTCTGAAATATCATA
TTTCTGTGTGTGTCTCTCTCAGCCCCCTGCCAGGCTAGAGGGAAACAGCTACTGATAATCGAA
AACTGCTGTTTGTGGCANGAACCCCTGGCTGTGCAAATAAATGGGGCTGAGGCCCTGTGTGA
TATTGAAGA

264/550

FIGURE 264

MGIQTSPVLLASLGVGLVTLLGLAVGSYLVRRSRRPQVTLLDPNEKYLLRLLDKTTVSHNTKR
FRFALPTAHHTLGLPVGKHIYLSTRIDGSLVIRPYTPVTSDEDQGYVDLVIKVYLKGVHPKFP
EGGKMSQYLDLKVGDVVEFRGPSGLLTYTGKGFNIQPNKKSPPPEPRVAKKLGMIAGGTGIT
PMLQLIRAILKVPEDPTQCFLLFANQTEKDILREDLEELQARYPNRFKLWFTLDHPPKDWAY
SKGFVTADMIREHLPAPGDDVLVLLCGPPPMVQLACHPNLDKLGYSQKMRFTY

Important features:

Signal peptide:

amino acids 1-26

N-glycosylation site:

amino acids 214-218

N-myristoylation sites:

amino acids 22-28, 76-82, 128-134, 180-186

265/550

FIGURE 265

CCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGCTTCGTTAGA
ACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACGATTTAGGTGACACTATAG
AATAACATCCACTTTGCCTTCTCTCCACAGGTGTCCACTCCCAGGTCCAAGTGCACCTCGGT
TCTATCGATAATCTCAGCACCAGCCACTCAGAGCAGGGCACG**ATG**TTGGGGGCCCCGCCTCAGG
CTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCATGAGCGTCCTCAGAGCCTATCCCAATGCC
TCCCCACTGCTCGGCTCCAGCTGGGGTGGCCTGATCCACCTGTACACAGCCACAGCCAGGAAC
AGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCATCAGACCATCTAC
AGTGCCCTGATGATCAGATCAGAGGATGCTGGCTTTGTGGTGATTACAGGTGTGATGAGCAGA
AGATACCTCTGCATGGATTTTCAGAGGCAACATTTTTGGATCACACTATTTTCGACCCGGAGAAC
TGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCCTCAGTATCAC
TTCCTGGTCAGTCTGGGCCGGGCGAAGAGAGCCTTCCTGCCAGGCATGAACCCACCCCCGTAC
TCCCAGTTCCTGTCCCGGAGGAACGAGATCCCCCTAATTCACTTCAACACCCCCATACCACGG
CGGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCTGAACGTGCTGAAGCCCCGG
GCCCCGATGACCCCGGCCCGGCCTCCTGTTACAGGAGCTCCCGAGCGCCGAGGACAACAGC
CCGATGGCCAGTGACCCATTAGGGGTGGTCAGGGGCGGTGAGTGAACACGCACGCTGGGGGA
ACGGGCCCCGAAGGCTGCCGCCCTTCGCCAAGTTCATC**TAG**GGTCGCTGG

266/550

FIGURE 266

MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSWSGGLIHLYTATARNSYHLQIHKNGHVDG
APHQTIYSALMIRSEDAGFVVITGVMSRRYLCMDFRGNIFGSHYFDPENCRFQHQTLENGYDV
YHSPQYHFLVSLGRAKRAFLPGMNPPYSQFLSRRNEIPLIHFNTPIPRRHTRSAEDDSERDP
LNVLKPRARMTPAPASCSQELPSAEDNSPMASDPLGVVRGGRVNT HAGGTGPEGCRPFAKFI

Important features:**Signal peptide:**

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 175-179

N-myristoylation site.

amino acids 33-39, 100-106, 225-231, 229-235

HBGF/FGF family proteins

amino acids 73-124

267/550

FIGURE 267

GGCTGAGGGGAGGCCCGAGCCTTTCTGGGGCCTGGGGGATCCTCTTGCACTGGTGGGTGGAGAGAAGCGCCTGC
AGCCAACCAGGGTCAGGCTGTGCTCACAGTTTCCTCTGGCGGCATGTAAAGGCTCCACAAAGGAGTTGGGAGTTC
AAATGAGGCTGTGCGGACGGCCTGAGGATGGACCCCAAGCCCTGGACCTGCCGAGCGTGGCACTGAGGCAGCGG
CTGACGCTACTGTGAGGGAAAGAAGGTTGTGAGCAGCCCCGAGGACCCCTGGCCAGCCCTGGCCCCAGCCTCTG
CCGGAGCCCTCTGTGGAGGCAGAGCCAGTGGAGCCCACTGAGGCAGGGCTGCTTGGCAGCCACCGGCCTGCAACT
CAGGAACCCCTCCAGAGGCCATGGACAGGCTGCCCCGCTGACGGCCAGGGTGAAGCATGTGAGGAGCCGCCCGG
AGCCAAGCAGGAGGGAAGAGGCTTTCATAGATTCTATTACAAAGAATAACCACCATTTTGCAAGGACCATGAGG
CCTACTGTGCGTGACATGCTGGTGGCTCGGACTGCTGGCTGCCATGGGAGCTGTTGCAGGCCAGGAGGACGGTTTT
GAGGGCACTGAGGAGGGCTCGCCAAGAGAGTTTACCTTAAACAGGTACAAGCGGGCGGGCGAGTCCCAGGAC
AAGTGCACCTACACCTTATTGTGCCCCAGCAGCGGGTCACGGGTGCCATCTGCGTCAACTCCAAGGAGCCTGAG
GTGCTTCTGGAGAACCGAGTGCATAAGCAGGAGCTAGAGCTGCTCAACAATGAGCTGCTCAAGCAGAAGCGGCAG
ATCGAGACGCTGCAGCAGCTGGTGGAGGTGGACGGCGGCATTGTGAGCGAGGTGAAGCTGCTGCGCAAGGAGAGC
CGCAACATGAACTCGCGGGTCACGCAGCTCTACATGCAGCTCCTGCACGAGATCATCCGCAAGCGGGACAACCGC
TTGGAGCTCTCCAGCTGGAGAACAGGATCCTGAACCAGACAGCCGACATGCTGCAGCTGGCCAGCAAGTACAAG
GACCTGGAGCACAAGTACCAGCACCTGGCCACACTGGCCCACAACCAATCAGAGATCATCGCGCAGCTTGAGGAG
CACTGCCAGAGGGTGCCCTCGGCCAGGCCCCGTCCCCAGCCACCCCCGCTGCCCCGCCCCGGGTCTACCAACCA
CCCACCTACAACCGCATCATCAACCAGATCTCTACCAACGAGATCCAGAGTGACCAGAACCTGAAGGTGCTGCCA
CCCCCTCTGCCCCTATGCCCCTCTCACCAGCCTCCCATCTTCCACCGACAAGCCGTGCGGGCCCATGGAGAGAC
TGCCTGCAGGCCCTGGAGGATGGCCACGACACAGCTCCATCTACCTGGTGAAGCCGGAGAACAACCAACCGCCTC
ATGCAGGTGTGGTGGGACAGAGACACGACCCCGGGGGCTGGACCGTCATCCAGAGACGCCTGGATGGCTCTGTT
AACTTCTTCAGGAAGTGGGAGACGTACAAGCAAGGGTTTGGGAACATTGACGGCGAATACTGGCTGGGCCTGGAG
AACATTTACTGGCTGACGAACCAAGGCAACTACAACTCCTGGTGACCATGGAGGACTGGTCCGGCCGCAAAGTC
TTTGCAGAATACGCCAGTTTCCGCCTGGAACCTGAGAGCGAGTATTATAAGCTGCGGCTGGGGCGCTACCATGGC
AATGCGGGTGACTCCTTTACATGGCACAACGGCAAGCAGTTACCCACCCTGGACAGAGATCATGATGTCTACACA
GGAACTGTGCCCCTACCAGAAGGGAGGCTGGTGGTATAACGCCTGTGCCCCTCCAACCTCAACGGGGTCTGG
TACCGCGGGGGCCATTACCGGAGCCGCTACCAGGACGGAGTCTACTGGGCTGAGTTCCGAGGAGGCTCTTACTCA
CTCAAGAAAGTGGTGATGATGATCCGACCGAACCCCAACACCTTCCACTTAAGCCAGCTCCCCCTCCTGACCTCTC
GTGGCCATTGCCAGGAGCCCACCCTGGTCACGCTGGCCACAGCACAAAGAACAACCTCCTCACCAGTTCATCCTGA
GGCTGGGAGGACCGGGATGCTGGATTCTGTTTTCCGAAGTCACTGCAGCGGATGATGGAAGTGAATCGATACGGT
GTTTTCTGTCCCTCCTACTTTCCTTCACACCAGACAGCCCCTCATGTCTCCAGGACAGGACAGGACTACAGACAA
CTCTTTCTTTAAATAAATTAAGTCTCTACAATAAAAAAAA

268/550

FIGURE 268

MRPLCVTCWWLGLLAAMGAVAGQEDGFEGTEEGSPREFIYLNRYKRAGESQDKCTYTFIVPQQ
RVTGAICVNSKEPEVLLNVRVHKQELELLNNELLKQKRQIETLQQLVEVDGGIVSEVKLLRKE
SRNMNSRVTQLYMQLLHEIIRKRDNALELSQLENRILNQTADMLQLASKYKDLEHKYQHLATL
AHNQSEIIAQLEEHQCQRVPSARFVPQPPPAAPPRVYQPPTYNRIINQISTNEIQSDQNLKVLP
PPLPTMPTLTSLPSSTDKPSGPWRDCLQALDGHDTSSIYLVKPENTNRLMQVWCDQRHDPGG
WTVIQRRLDGSVNFFRNWETKQGFNGIDGEYWLGLENIYWLTNQGNKLLVTMEDWSGRKVF
AEYASFRLEPESEYYKLRLGRYHGNAGDSFTWHNGKQFTTLDRDHDVYTGNCAYQKGGWWYN
ACAHSNLNGVWYRGGHYRSRYQDGVYWAEFRGGSYSLKKVMMIRPNPNTFH

Important features:**Signal peptide:**

amino acids 1-22

N-glycosylation sites:

amino acids 164-168, 192-196

cAMP- and cGMP-dependent protein kinase phosphorylation site:

amino acids 124-128

Tyrosine kinase phosphorylation sites:

amino acids 177-184, 385-393, 385-394, 461-468

N-myristoylation sites:amino acids 12-18, 18-24, 22-28, 29-35, 114-120, 341-347, 465-471,
473-479**Amidation site:**

amino acids 373-377

Fibrinogen beta and gamma chains C-terminal domain signature:

amino acids 438-451

Fibrinogen beta and gamma chains C-terminal domain proteins:

amino acids 305-343, 365-402, 411-424, 428-458

Trehalase proteins:

amino acids 275-292

269/550

FIGURE 269

GCCGAGCTGAGCGGATCCTCACATGACTGTGATCCGATTCTTTCCAGCGGCTTCTGCAACCAA
GCGGGTCTTACCCCCGGTCTCCGCGTCTCCAGTCTTCGCACCTGGAACCCCAACGTCCCCGA
GAGTCCCCGAATCCCCGCTCCCAGGCTACCTAAGAGGATGAGCGGTGCTCCGACGGCCGGGGC
AGCCCTGATGCTCTGCGCCGCCACCGCCGTGCTACTGAGCGCTCAGGGCGGACCCGTGCAGTC
CAAGTCGCCGCGCTTTGCGTCTCTGGGACGAGATGAATGTCTGGCGCACGGACTCCTGCAGCT
CGGCCAGGGGCTGCGCGAACACGCGGAGCGCACCCGAGTCAGCTGAGCGCGCTGGAGCGGCG
CCTGAGCGCGTGCGGGTCCGCTGTGAGGGAACCGAGGGGTCCACCGACCTCCCGTTAGCCCC
TGAGAGCCGGGTGGACCCTGAGGTCTTCACAGCCTGCAGACACAACCTCAAGGCTCAGAACAG
CAGGATCCAGCAACTCTTCCACAAGGTGGCCAGCAGCAGCGGCACCTGGAGAAGCAGCACCT
GCGAATTCAGCATCTGCAAAGCCAGTTTGGCCTCTGGACCACAAGCACCTAGACCATGAGGT
GGCCAAGCCTGCCCCAAGAAAGAGGCTGCCCGAGATGGCCCAGCCAGTTGACCCGGCTCACAA
TGTCAGCCGCTGCACCGGCTGCCAGGGATTGCCAGGAGCTGTTCCAGGTGGGGAGAGGCA
GAGTGGACTATTTGAAATCCAGCCTCAGGGGTCTCCGCCATTTTGGTGAAGTCAAGATGAC
CTCAGATGGAGCTGGACAGTAATTCAGAGGCGCCACGATGGCTCAGTGGACTTCAACCGGCC
CTGGGAAGCCTACAAGCGGGGTTTGGGGATCCCCACGGCGAGTTCTGGCTGGGTCTGGAGAA
GGTGCATAGCATCACGGGGGACCGCAACAGCCGCTGGCCGTGCAGCTGCGGGACTGGGATGG
CAACGCCGAGTTGCTGCAGTTCTCCGTGCACCTGGGTGGCGAGGACACGGCCTATAGCCTGCA
GCTCACTGCACCCGTGGCCGGCCAGCTGGGCGCCACCACCGTCCCACCCAGCGGCCTCTCCGT
ACCCCTTCTCCACTTGGGACCAGGATCACGACCTCCGCAGGGACAAGAACTGCGCCAAGAGCCT
CTCTGGAGGCTGGTGGTTTGGCACCTGCAGCCATTCCAACCTCAACGGCCAGTACTTCCGCTC
CATCCACAGCAGCGGCAGAAGCTTAAGAAGGGAATCTTCTGGAAGACCTGGCGGGGCCGCTA
CTACCCGCTGCAGGCCACCACCATGTTGATCCAGCCCATGGCAGCAGAGGCAGCCTCCTAGCG
TCCTGGCTGGGCCTGGTCCCAGGCCACGAAAGACGGTGAAGTCTTGGCTCTGCCCAGGATGT
GGCCGTTCCCTGCCTGGGCAGGGGCTCCAAGGAGGGGCCATCTGGAACTTGTGGACAGAGAA
GAAGACCAGACTGGAGAAGCCCCCTTTCTGAGTGCAGGGGGGCTGCATGCGTTGCCTCCTGA
GATCGAGGCTGCAGGATATGCTCAGACTCTAGAGGCGTGGACCAAGGGGCATGGAGCTTCACT
CCTTGCTGGCCAGGGAGTTGGGACTCAGAGGGACCACTTGGGGCCAGCCAGACTGGCCTCAA
TGGCGGACTCAGTCACATTGACTGACGGGGACAGGGCTTGTGTGGGTCGAGAGCGCCCTCAT
GGTGCTGGTGCTGTTGTGTGTAGGTCCCCTGGGGACACAAGCAGGCGCCAATGGTATCTGGG
GGAGCTCACAGAGTTCTTGAATAAAAGCAACCTCAGAACAC

270/550

FIGURE 270

MTVIRFFPAASATKRVLPVLRVSSPRTWNPVPESPRIPAPRLPKRMSGAPTAGAALMLCAA
TAVLLSAQGGPVQSKSPRFASWDEMNVLAHGLLQLGQGLREHAERTRSQLSALERRLSACGSA
CQGTEGSTDLPLAPESRVDPEVLHSLQTQLKAQNSRIQQLFHKVAQQQRHLEKQHLRIQHLS
QFGLLDHKHLDHEVAKPARRKRLPEMAQPVDPAHNVSRHLHRLPRDCQELFQVGERQSGLF
PQGSPPFLVNCKMTSDGGWTVIQRRHDGSDVFNRPWEAYKAGFGDPHGEFWLGLEKVHSITGD
RNSRLAVQLRDWDGNAELLQFSVHLGGEDTAYSLQLTAPVAGQLGATTVPSPGLSVPFSTWDQ
DHDLRDKNCAKSLSGGWVFGTCSHSLNNGQYFRSIPQQRQKLKKGIFWKTWRGRYYPLQATT
MLIQPMAAEAAAS

Important features:**Signal peptide:**

Amino acids 1-13

Transmembrane domain:

Amino acids 53-70

N-glycosylation site:

Amino acids 224-228

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 46-50;118-122

N-myristoylation sites:

Amino acids 50-56;129-135;341-347;357-363

Fibrinogen beta and gamma chains C-terminal domain signature:

Amino acids 396-409

271/550

FIGURE 271

CGGACGCGTG GGGGAAACCCCTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGGG
AACAAGATGGCGGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAAGCTGGGGCTCCCGCCGCTG
CTGCTGCTGACCATGGCCTTGCCCGGAGGTTCCGGGACCGCTTCGGCTGAAGCATTGTGACTCG
GTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCACACCTAC
CCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTCAGTTT
GTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA
TATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCCATTCGCT
GAACTGAGACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTCCTCTAACT
CTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACCTCTTCATGG
ACTTTTTATCTTCAAGCCGATGACGGAAAAATAGTTATATTCCAGTCTAAGCCAGAAATCCAG
TACGCACCACATTTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAAGCAAAATGTCC
TATCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTTCTTGAAGATGGAGAAAGTGATGGC
TTTTTAAGATGCCTCTCTCTTAACCTCTGGGTGGATTTTAACTACAACCTCTTGTCTCTCGGTG
ATGGTATTGCTTTGGATTTGTTGTGCAACTGTTGCTACAGCTGTGGAGCAGTATGTTCCCTCT
GAGAAGCTGAGTATCTATGGTGACTTGGAGTTTATGAATGAACAAAAGCTAAACAGATATCCA
GCTTCTTCTCTTGTGGTTGTTAGATCTAAACTGAAGATCATGAAGAAGCAGGGCCTCTACCT
ACAAAAGTGAATCTTGCTCATTCTGAAATTTAAGCATTCTTTTAAAGACAAGTGTAATA
GACATCTAAAATTCACCTCCTCATAGAGCTTTTAAATGGTTTCATTGGATATAGGCCTTAAG
AAATCACTATAAAATGCAAATAAAGTTACTCAAATCTGTG

272/550

FIGURE 272

MAAPKGSLWVRTQLGLPPLLLLTMALAGGSGTASAEAFDSVLGDTASCHRAQCLTYPLHTYPK
EEELYACQRGCRFLFSICQFVDDGIDLNRKLECESACTEAYSQSDEQYACHLGCONQLPFAEL
RQEQLMSLMPKMHLLFPLTLVRSFWSMMDSAQSFITSSWTFYQLQADDGKIVIFQSKPEIQYA
PHLEQEPTNLRESSLSKMSYLMRNSQAHRNFLEDGESDGFLRCLSLNSGWILTTTLVLSVMV
LLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAGPLPTK
VNLAHSEI

Important features:**Signal peptide:**

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site:

amino acids 90-94

N-myristoylation sites:

amino acids 28-34, 29-35, 31-37, 86-92

273/550

FIGURE 273

CCCACGCGTCCGAACCTCTCCAGCGATGGGAGCCGCCCGCCTGCTGCCCAACCTCACTCTGTG
CTTACAGCTGCTGATTCTCTGCTGTCAAACCTCAGTACGTGAGGGACCAGGGCGCCATGACCGA
CCAGCTGAGCAGGCGGCAGATCCGCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCACGT
GCAGGTACCGGGCGTCGCATCTCCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAGCTCAT
AGTGGAGACGGACACGTTTGGCAGCCGGGTTTCGCATCAAAGGGGCTGAGAGTGAGAAGTACAT
CTGTATGAACAAGAGGGGCAAGCTCATCGGGAAGCCCAGCGGGAAGAGCAAAGACTGCGTGTT
CACGGAGATCGTGCTGGAGAACAACCTATACGGCCTTCCAGAACGCCCGGCACGAGGGCTGGTT
CATGGCCTTCACGCGGCAGGGGCGGCCCCGCCAGGCTTCCCGCAGCCGCCAGAACCAGCGCGA
GGCCCACTTCATCAAGCGCCTCTACCAAGGCCAGCTGCCCTTCCCCAACCACGCCGAGAAGCA
GAAGCAGTTCGAGTTTGTGGGCTCCGCCCCCACC CGCGGACCAAGCGCACACGGCGGCCCCA
GCCCCTCACGTAGTCTGGGAGGCAGGGGGCAGCAGCCCCCTGGGCCGCCTCCCCACCCCTTTCC
CTTCTTAATCCAAGGACTGGGCTGGGGTGGCGGGAGGGGAGCCAGATCCCCGAGGGAGGACCC
TGAGGGCCGCGAAGCATCCGAGCCCCCAGCTGGGAAGGGGCAGGCCGGTGCCCCAGGGGCGGC
TGGCACAGTGCCCCCTTCCCGGACGGGTGGCAGGCCCTGGAGAGGAACTGAGTGTCACCCTGA
TCTCAGGCCACCAGCCTCTGCCGGCCTCCCAGCCGGGCTCCTGAAGCCCGCTGAAAGGTCAGC
GACTGAAGGCCTTGACAGACAACCGTCTGGAGGTGGCTGTCCTCAAAATCTGCTTCTCGGATCT
CCCTCAGTCTGCCCCCAGCCCCCAAACCTCCTCCTGGCTAGACTGTAGGAAGGGACTTTTGTTT
GTTTGTTTGTTCAGGAAAAAAGAAAGGGAGAGAGAGGAAAATAGAGGGTTGTCCACTCCTCA
CATTCCACGACCCAGGCCTGCACCCCCACCCCCAACTCCCAGCCCCGGAATAAAACCATTTTCC
TGC

274/550

FIGURE 274

MGAARLLPNLTLCLQLLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRIS
ATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLGKPSGKSKDCVFTEIVLENN
YTAFQONARHEGWFMATFTRQGRPRQASRSRQNRQEAHFIRLYQGQLPFPNHAEKQKQFEFVGS
APTRRTKRTRRPQPLT

Important features:

Signal peptide:

Amino acids 1-22

N-glycosylation site.

amino acids 9-13, 126-130

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 60-64

Casein kinase II phosphorylation site.

amino acids 65-69

Tyrosine kinase phosphorylation site.

amino acids 39-48, 89-97

N-myristoylation site.

amino acids 69-75, 188-194

Amidation site.

amino acids 58-62

HBGF/FGF family signature.

amino acids 103-128

275/550

FIGURE 275

TATTTACCATATCAGATTCACATTCAGTCCTCAGCAAAATGAAGGGCTCCATTTTCACTCTGT
TTTTATTCTCTGTCCTATTTGCCATCTCAGAAGTGCGGAGCAAGGAGTCTGTGAGACTCTGTG
GGCTAGAATACATACGGACAGTCATCTATATCTGTGCTAGCTCCAGGTGGAGAAGGCATCTGG
AGGGGATCCCTCAAGCTCAGCAAGCTGAGACAGGAACTCCTTCCAGCTCCCACATAAACGTG
AGTTTTCTGAGGAAAATCCAGCGCAAAACCTTCCGAAGGTGGATGCCTCAGGGGAAGACCGTC
TTTGGGGTGGACAGATGCCCCACTGAAGAGCTTTGGAAGTCAAAGAAGCATTCAAGTGATGTCAA
GACAAGATTTACAACTTTGTGTTGCACTGATGGCTGTTCCATGACTGATTTGAGTGCTCTTT
GCTAAAGACAAGAGCAAATACCCAATGGGTGGCAGAGCTTTATCACATGTTTAATTACAGTGTT
TTACTGCCTGGTAGAACACTAATATTGTGTTATTAAAATGATGGCTTTTGGGTAGGCAAACT
TCTTTTCTAAAAGGTATAGCTGAGCGGTGAAACCACAGTGATCTCTATTTTCTCCCTTTGCC
AAGGTTAATGAACTGTTCTTTTCAAATTCTACTAATGCTTTGAAATTTCAAATGCTGCGCAAA
ATTGCAATAAAAATGCTATAAA

276/550

FIGURE 276

MKGSIFTLFLFSVLFAISEVRSKESVRLCGLEYIRTVIYICASSRWRRLHLEGIPQAQQAETGN
SFQLPHKREFSEENPAQNLPKVDASGEDRLWGGQMPTEELWKSKKHSVMSRQDLQTLCTDGC
SMTDLSALC

Important features:

Signal sequence:

amino acids 1-18

cAMP- and cGMP-dependent protein kinase phosphorylation site:

amino acids 107-111

N-myristoylation sites:

amino acids 3-9, 52-58, 96-102, 125-131

Insulin family signature:

amino acids 121-136

Insulin family proteins:

amino acids 28-46

277/550

FIGURE 277

GCAGCTGGTTACTGCATTTCTCCATGTGGCAGACAGAGCAAAGCCACAACGCTTTCTCTGCTGGATTAAAGACGG
CCCACAGACCAGAACTTCCACTATACTACTTAAATACATAGGTGGCTTGTCAAATTCATTGATTAGTATTGT
AAAAGGAAAAAGAAGTTCTTCTTACAGCTTGGATTCAACGGTCCAAAACAAAATGCAGCTGCCATTAAAGTCT
CAGATGAACAAACTTCTACACTGATTTTAAATCAAGAATAAGGGCAGCAAGTTTCTGGATTCACTGAATCAAC
AGACACAAAAAGCTGGCAATATAGCAACTATGAAGAGAAAAGCTACTAATAAAATTAACCCACGCATAGAAGAC
TTTTTTTTCTCTCTTCTAAAAACAATAAGTAAAGACTTAAATTTAAACACATCATTTTACAACCTCATTTCAAAAT
GAAGACTTTTACCTGGACCCTAGGTGTGCTATTCTTCCTACTAGTGGACACTGGACATTGCAGAGGTGGACAATT
CAAAATTAATAAAATAAAACCAGAGAAGATACCCTCGTGGCCACAGATGGTAAAGAGGAAGCAAAGAAATGTGCATA
CACATTCCTGGTACCTGAACAAAGAATAACAGGGCCAATCTGTGTCAACACCAAGGGGCAAGATGCAAGTACCAT
TAAAGACATGATCACCAGGATGGACCTTGAAAACCTGAAGGATGTGCTCTCCAGGCAGAAGCGGGAGATAGATGT
TCTGCAACTGGTGGTGGATGTAGATGGAAACATTGTGAATGAGGTAAAGCTGCTGAGAAAGGAAAGCCGTAACAT
GAACCTCTCGTGTACTCAACTCTATATGCAATTATTACATGAGATTATCCGTAAGAGGGATAATTCACTTGAAT
TTCCCAACTGGAAAACAAAATCCTCAATGTCACCACAGAAAATGTTGAAGATGGCAACAAGATACAGGGAAC TAGA
GGTGAAATACGCTTCTTGACTGATCTGTCAATAACCAATCTGTGATGATCACTTTGTTGGAAGAACAGTGCCT
GAGGATATTTTCCCGACAAGACACCCATGTGTCTCCCCACTTGTCCAGGTGGTGGCCACAACATATTCCTAACAG
CCAACAGTATACTCCTGGTCTGCTGGGAGGTAACGAGATTGAGAGGGATCCAGGTTATCCAGAGATTTAATGCC
ACCACCTGATCTGGCAACTTCTCCACCAAAGCCCTTTCAAGATACCACCGGTAACCTTCATCAATGAAGGACC
ATTCAAAGACTGTCAGCAAGCAAAGAAGCTGGGCATTTCGGTCAGTGGGATTTATATGATTAAACCTGAAAACAG
CAATGGACCAATGCAGTTATGGTGTGAAAACAGTTTGGACCCTGGGGTTGGACTGTTATTGAGAAAAGAACAGA
CGGCTCTGTCAACTTCTTCAGAAATTTGGGAAAATTTAAGAAAGGGTTTGGAAACATTGACGGAGAATACTGGCT
TGGACTGGAAAATATCTATATGCTTAGCAATCAAGATAATTACAAGTTATTGATTGAATTAGAAGACTGGAGTGA
TAAAAAAGTCTATGCAGAATACAGCAGCTTTTCGTCTGGAACCTGAAAGTGAATTCTATAGACTGCGCCTGGGAAC
TTACCAGGGAATGCAGGGGATTCTATGATGTGGCATAATGGTAAACAATTCACCACACTGGACAGAGATAAAGA
TATGTATGCAGGAACTGCGCCACTTTCATAAAGGAGGCTGGTGGTACAATGCCTGTGCACATTCTAACCTAAA
TGGAGTATGGTACAGAGGAGGCCATTACAGAAGCAGCACCAAGATGGAATTTTCTGGGCCGAATACAGAGGCGG
GTCATACTCCTTAAGAGCAGTTGAGATGATGATCAAGCCTATTGACTGAAGAGAGACACTCGCCAATTTAAATGA
CACAGAACTTTGTACTTTTCAGCTCTTAAAAATGTAAATGTTACATGTATATTACTTGGCACAATTTATTTCTAC
ACAGAAAGTTTTTAAATGAATTTTACCGTAACATAAAAGGGAACCTATAAATGTAGTTTCATCTGTCTGCTCAAT
TACTGCAGAAAATTATGTGTATCCACAACCTAGTTATTTTAAAAATTATGTTGACTAAATACAAAGTTTGTTC
TAAATGTAAATATTTGCCACAATGTAAAGCAAATCTTAGCTATATTTTAAATCATAAATAACATGTTCAAGATA
CTTAACAATTTATTTAAATCTAAGATTGCTCTAACGTCTAGTGAAAAAAATATTTTTTAAATTTTACGCCAAATA
ATGCATTTTATTTTATAAAAAATACAGACAGAAAATTAGGGAGAACTTCTAGTTTTGCCAATAGAAAATGTTCTT
CCATTGAATAAAAGTTATTTCAAATTGAATTTGTGCCTTTCACACGTAATGATTAAATCTGAATTTCTAATAATA
TATCCTATGCTGATTTTCCAAAACATGACCCATAGTATTAATAACATATCATTTTTTAAAAATAAAAAAAACCC
AAAAATAATGCATGCATAATTTAAATGGTCAATTTATAAAGACAAATCTATGAATGAATTTTTCAGTGTATCTT
CATATGATATGCTGAACACCAAATCTCCAGAAATGCATTTTATGTAGTTCTAAAAATCAGCAAAATATTGGTATT
ACAAAATGCAGAATATTTAGTGTGCTACAGATCTGAATTATAGTTCTAATTTATTATTACTTTTTTCTAATTT
ACTGATCTTACTACTACAAAGAAAAAAACCCAACCCATCTGCAATTCAAATCAGAAAAGTTTGGACAGCTTTAC
AAGTATTAGTGCATGCTCAGAACAGGTGGGACTAAAACAACTCAAGGAAGTGTGGCTGTTTTCCCGATACTGA
GAATTCACAGCTCCAGAGCAGAAGCCACAGGGGCATAGCTTAGTCCAACTGCTAATTTTCAATTTTACAGTGTAT
GTAACGCTTAGTCTCACAGTGTCTTAACTCATCTTTGCAATCAACAACCTTACTAGTGACTTTCTGGAACAATT
TCCTTTCAGGAATACATATTCAGTCTTAGAGGTGACCTTGCCCTTAATATATTTGTGAAGTTAAATTTTAAAGA
TAGCTCATGAACTTTTGCTTAAGCAAAAAGAAAACCTCGAATTGAAATGTGTGAGGCAAACTATGCATGGGAAT
AGCTTAATGTGAAGATAATCATTTGGACAACCTCAAATCCATCAACATGACCAATGTTTTTCATCTGCCACATCTC
AAAATAAACTTCTGGTGAACAAATTAACAAAATATCCAAACCTCAAAAAAA

278/550

FIGURE 278

MKTFTWTLGVLEFLLVDTGHCRGGQFKIKKINQRRYPRATDGKEEAKKCAYTFLVPEQRITGP
ICVNTKGQDASTIKDMITRMDLENLKDVLRSQKREIDVLQLVVDVDGNIVNEVKLLRKESRNM
NSRVTQLYMQLLHEIIRKRDNSLELSQLENKILNVTTEMLKMATRYRELEVKYASLTDLVNNQ
SVMITLLEEQCLRIFSRQDTHVSPPLVQVVPQHIPNSQQYTPGLLGGNEIQRDPGYPRDLMPP
PDLATSPTKSPFKIPPVTFINEGPFKDCQQAKEAGHSVSGIYMIKPENSNGPMQLWCENSLDP
GGWTVIQKRTDGSVNFFRNWENYKKGFGNIDGEYWLGLENIYMLSNQDNYKLLIELEDWSDKK
VYAEYSSFRLEPESEFYRLRLGTYQGNAGDSMMWHNGKQFTTLDRDKDMYAGNCAHFHKGWW
YNACAHNSNLNGVWYRGGHYRSKHQDGIFWAEYRGGSYSLRAVQMMIKPID

Important features:**Signal sequence:**

Amino acids 1-23

N-glycosylation sites:

Amino acids 160-164;188-192

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 120-124

Tyrosine kinase phosphorylation sites:

Amino acids 173-180;387-396

N-myristoylation sites:Amino acids 70-76;110-116;232-238,343-349;400-406;467-473;
475-487**Fibrinogen beta and gamma chains C-terminal domain signature:**

Amino acids 440-453

279/550

FIGURE 279

CCCACGCGTCCGCGCAGTCGCGCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCGGC
CCGGGGCTGTGGCGTCGACTCCGACCCAGGCAGCCAGCAGCCCGCGCGGGAGCCGGACCGCCG
CCGGAGGAGCTCGGACGGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGC GGAGAAGCC
CGGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGAGACAGCGGACAAGCAGCGGA
GGAGAAGGAGGAGGAGGCGAACCCAGAGAGGGGCGAGCAAAGAAGCGGTGGTGGTGGGCGTCG
TGGCC**ATG**GCGGCGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCGCGAGCGCG
AGAAATCCAACGCCTGCAAGTGTGTCAGCAGCCCCAGCAAAGGCAAGACCAGCTGCGACAAAA
ACAAGTTAAATGTCTTTTCCCGGGTCAAACCTCTTCGGCTCCAAGAAGAGGGCGCAGAAGAAGAC
CAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACCACTTGCAGC
TGCAGGCGGATGGAACCATTTGATGGCACCAAAGATGAGGACAGCACTTACACTCTGTTTAACC
TCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCTGTACTTGGCAA
TGAACAGTGAGGGATACTTGTACACCTCGGAACCTTTTCACACCTGAGTGCAAATTCAAAGAAT
CAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAGCAGCAGTCAGGCC
GAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAACCATGTGAAGAAGA
ACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGTACAAGGAGCCATCAC
TGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACCAAGAGCAGAAGTGTCT
CTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAACG**TAG**CCAGTGAGGGCAA
AAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAATTCTTCTAGCAGTCCTTCA
CCCAAAGTTCAAATTTGTCAGTGACATTTACCAAACAAACAGGCAGAGTTCACTATTCTATC
TGCCATTAGACCTTCTTATCATCCATACTAAAGC

280/550

FIGURE 280

MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCDKNKLNVSFVKLFGSKKRRRRRPE
PQLKGIVTKLYSRQGYHLQLQADGTIDGTDKEDSTYTLFNLI PVGLRVVAIQGVQTKLYLAMN
SEGILYTSELFTPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLNKEGEIMKGNHVKKNK
PAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKSRSVSGVLNGGKSMHNEST

Important Features:

N-glycosylation site:

Amino acids 242-246

Glycosaminoglycan attachment sites:

Amino acids 165-169, 218-222

Tyrosine kinase phosphorylation site:

Amino acids 93-100

N-myristoylation sites:

Amino acids 87-93, 231-237

ATP/GTP-binding site motif A (P-loop):

Amino acids 231-239

HBGF/FGF family proteins:

Amino acids 78-94, 102-153

281/550

FIGURE 281

CCAGGATGGAGCTGGGGCCTGTATAGCCATATTATTGTTCTATGCTACTAGACATGGGGGGGA
CTTGGTGAAAAAGGTATTATCCAGCCAGAGGGTCTGGGAGCCCTGTCTTACTGAACCTGGGCA
ACCTGGATATTCTGAGACATATTTTGGGGGGATTTCAGTGAAAAAGTGGGGGATCCCCCTCCA
TTTAGAGTGTAGCAAAGGAAAAACACCAAGGTTGGGTTCCCTCCTGACATTGGCAGTGCCCC
AGTAGGGGTGGGATGAGCGAATATTTCCCAAAGCTAAAGTCCACACCCTGTAGATTACAAGAG
TGGATTTGGCAGGAGTGTGCCCCAAAATACAGTGGAAGGTGCCTGAAGATATTTAAACCACG
TCTTGGAATTTAGTGGGTCTTGGCTTTGGGATAGGTGAAGTGAGGACAGACACTGGAGAGGA
GGGAAAGGGGACGTTTTCAATAGGAGGCCAAACTCGAGGGTGGGATCCACTGAGGAGTACATA
GGCTGCTGGATCTGGTGGAGCCAGCACTGGGCCCACGGGTGGTAACTGGCTGCTGTGGAGGGG
GGTACGTGAGGGGGGGGTCTGGGGCTTATCCTCAGGTCTGTGGGTGGGGCAGCGAGTCGGGG
CCTGAGCGTCAAGAGCATGCCCTAGTGAGCGGGCTCCTCTGGGGGAGCCCAGCGCGCTCCGGG
CGCCTGCCGGTTTGGGGGTGTCTCCTCCCGGGGCGCT~~AT~~GGCGGCGCTGGCCAGTAGCCTGAT
CCGGCAGAAGCGGGAGGTCCGCGAGCCCGGGGGCAGCCGGCCGGTGTGCGCGCAGCGGCGCGT
GTGTCCCCCGCGCACCAAGTCCCTTTGCCAGAAGCAGCTCCTCATCCTGCTGTCCAAGGTGCG
ACTGTGCGGGGGGCGGCCCCGCGCGGCCGACCGCGGCCCGGAGCCTCAGCTCAAAGGCATCGT
CACCAAAGTGTCTGCCGCCAGGGTTTCTACCTCCAGGCGAATCCCGACGGAAGCATCCAGGG
CACCCCAGAGGATACCAGCTCCTTCACCCACTTCAACCTGATCCCTGTGGGCCTCCGTGTGGT
CACCATCCAGAGCGCCAAGCTGGGTCACTACATGGCCATGAATGCTGAGGGACTGCTCTACAG
TTCGCCGCATTTACAGCTGAGTGTGCTTTAAGGAGTGTGTCTTTGAGAATTACTACGTCCT
GTACGCCTCTGCTCTCTACCGCCAGCGTCGTTCTGGCCGGGCCTGGTACCTCGGCCTGGACAA
GGAGGGCCAGGTCATGAAGGGAAACCGAGTTAAGAAGACCAAGGCAGCTGCCCCACTTTCTGCC
CAAGCTCCTGGAGGTGGCCATGTACCAGGAGCCTTCTCTCCACAGTGTCCCCGAGGCCTCCCC
TTCCAGTCCCCCTGCCCCCTGA~~AA~~ATGTAGTCCCTGGACTGGAGGTTCCCTGCACTCCCAGTGA
GCCAGCCACCACCACAACCTGT

282/550

FIGURE 282

MAALASSLIRQKREVREPGGSRPVSAQRRVCPRGTKSLCQKQLLILLSKVRLCGGRPARPDRG
PEPQLKGIVTKLFCRQGFIYLAQNPDGSIQGTPEDTSSFTFHNLI PVGLRVVTIQSAKLGHYMA
MNAEGLLYSSPHFTAECRFKECVFENYYVLYASALYRQRRSGRAWYLGLDKEGQVMKGNRVKK
TKAAAHFLPKLLEVAMYQEPSLHSVPEASPSPPAP

Important features:

Tyrosine kinase phosphorylation site:

Amino acids 199-207

N-myristoylation sites:

Amino acids 54-60; 89-95; 131-137

HBGF/FGF family signature:

Amino acids 131-155

283/550

FIGURE 283

ATGCCCGCGGCCATCGCTAGCGGCTTGATCCGCCAGAAGCGGCAGGCGCGGGAGCAGCACTGG
GACCGGCCGTCTGCCAGCAGGAGGCGGAGCAGCCCCAGCAAGAACCGCGGGCTCTGCAACGGC
AACCTGGTGGATATCTTCTCCAAAGTGCGCATCTTCGGCCTCAAGAAGCGCAGGTTGCGGCGC
CAAGATCCCCAGCTCAAGGGTATAGTGACCAGGTTATATTGCAGGCAAGGCTACTACTTGCAA
ATGCACCCCGATGGAGCTCTCGATGGAACCAAGGATGACAGCACTAATTCTACACTCTTCAAC
CTCATACCAGTGGGACTACGTGTTGTTGCCATCCAGGGAGTGAAAACAGGGTTGTATATAGCC
ATGAATGGAGAAGGTTACCTCTACCCATCAGAACTTTTTACCCCTGAATGCAAGTTTAAAGAA
TCTGTTTTTGAATAATTATTATGTAATCTACTCATCCATGTTGTACAGACAACAGGAATCTGGT
AGAGCCTGGTTTTTGGGATTAAATAAGGAAGGGCAAGCTATGAAAGGGAACAGAGTAAAGAAA
ACCAAACCAGCAGCTCATTTTCTACCCAAGCCATTGGAAGTTGCCATGTACCGAGAACCATCT
TTGCATGATGTTGGGGAAACGGTCCCGAAGCCTGGGGTGACGCCAAGTAAAAGCACAAGTGCG
TCTGCAATAATGAATGGAGGCAAACCAGTCAACAAGAGTAAGACAACA**TAG**

284/550

FIGURE 284

MAAAIASGLIRQKRQAREQHWDRPSASRRRSSPSKNRGLCNGNLVDIFSKVRI FGLKKRRLRR
QDPQLKGIVTRLYCRQGYLQMHPDGALDGTKDDSTNSTLFNLI PVGLRVVAIQGVKTGLYIA
MNGEGYLYPSELF TPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNRVKK
TKPAAHFLPKPLEVAMYREPSLHDVGETVPKPGVTPSKSTSASAIMNGGKPVNKSITT

Important features:**N-glycosylation sites:**

Amino acids 100-104, 242-246

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 28-32, 29-33

Tyrosine kinase phosphorylation site:

Amino acids 199-207

N-myristoylation sites:

Amino acids 38-44, 89-95, 118-124, 122-128, 222-228

HBGF/FGF family proteins:

Amino acids 104-155, 171-198

285/550

FIGURE 285

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTCAGGTCCAGGTTTTGCTTTGA
TCCTTTTCAAAAAGTGGAGACACAGAAGAGGGCTCTAGGAAAAAGTTTTGGATGGGATTATGTGGAAGTACCCT
GCGATTCTCTGCTGCCAGAGCAGGCTCGGCGCTTCCACCCAGTGCAGCCTTCCCCTGGCGGTGGTGAAAGAGAC
TCGGGAGTCGCTGCTTCCAAAGTCCCCGCCGTGAGTGAGCTCTCACCCAGTGCAGCCAAATGAGCCTCTTCGGGC
TTCTCCTGCTGACATCTGCCCTGGCCGGCCAGAGACAGGGGACTCAGGCGGAATCCAACCTGAGTAGTAAATTCC
AGTTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAATTATTACTGTGTCTACTAATG
GAAGTATTCACAGCCCAAGGTTTCCTCATACTTATCCAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAG
AGGAAAATGTATGGATACAACTTACGTTTGATGAAAGATTTGGGCTTGAAGACCCAGAAGATGACATATGCAAGT
ATGATTTTGTAGAAGTTGAGGAACCCAGTGATGGAACCTATATTAGGGCGCTGGTGTGGTTCTGGTACTGTACCAG
GAAAACAGATTTCTAAAGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCCTTCTGAACAGGGT
TCTGCATCCACTACAACATTGTCTATGCCACAATTACAGAAGCTGTGAGTCCTTCAGTGTACCCCCCTTCAGCTT
TGCCACTGGACCTGCTTAATAATGCTATAACTGCCTTTAGTACCTTGAAGACCTTATTCGATATCTTGAACCAG
AGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTTCTTGGCAAGGCTTTTGTGTTTTGGAA
GAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACCTTCT
CAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTG
GTGGGAACTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACC
ACGAGGTCTTCAGTTGAGACCAAAGACCGGTGTGAGGGGATTGCACAAATCACTACCGACGTGGCCCTGGAGC
ACCATGAGGAGTGTGACTGTGTGTGTCAGAGGGAGCACAGGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCCA
GAGCTGTGAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGCT
TCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCA
ACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAATGTTGTAT
TAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTTCAGTCTTTTC
GATACGGCTTAGGGTAATGTGCTACAGGAAAAAACTGTGCAAGTGAACCTGATTCCGTTGCCTTGCTTAAC
TCTAAAGCTCCATGTCTGGGCCATAAATCGTATAAATCTGGATTTTTTTTTTTTTTTTTTGTCTCATATTCACAT
ATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACATATGTTGCTATGAATTAACTTGT
GTCATGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAATTTCTGCCATTTAGAAGAAGAGAACTACA
TTCATGGTTTGGAAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTG
TTTCATTGTGTACATTTTTATATCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTTAAATATATCT
ATTTTTACCAAAGGTATTTAATATTCTTTTTATGACAACTTAGATCAACTATTTTGTAGCTTGGTAAATTTTCT
AAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAATATTGTTGCTCTGACAAAAATACATGTATTTCA
TTCTCGTATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAA
GACTTTTTGAAAATAATTAATATATCTTCCATTCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGA
AAGTAGACATTCAGATCCAGCCATTACTAACCTATTCTTTTTTGGGGAATCTGAGCCTAGCTCAGAAAAACAT
AAAGCACCTTGAAAAGACTTGGCAGCTTCTGATAAAGCGTGCTGTGCTGTGCAGTAGGAACACATCCTATTTA
TTGTGATGTTGTGGTTTTTATTATCTTAACTCTGTTCCATACACTTGTATAAATACATGGATATTTTATGTACA
GAAGTATGTCTCTTAACAGTTCACTTATTGTACTCTGGCAATTTAAAAGAAAATCAGTAAATATTTTGTCTGT
AAAATGCTTAATATNGTGCCTAGGTTATGTGGTGACTATTTGAATCAAAAATGTATTGAATCATCAAATAAAGA
ATGTGGCTATTTTGGGAGAAAATTAATAAAAAAAAAAAAAAAAAAAGGTTTAGGGATAACAGGGTAATGCGGCC

286/550

FIGURE 286

MSLFGLLLLSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPRF
PHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWCGS
GTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLLNNA
ITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNLLEEVRLYSCTP
RNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVPSKVTKKYHEVLQLRPKT
GVRGLHKSLTDVALEHHEECDVCVRGSGTGG

Important features:**signal sequence:**

Amino acids 1-14

N-glycosylation sites:

Amino acids 25-29;55-59;254-258

N-myristoylation sites:

Amino acids 15-21;117-123;127-133;281-287;282-288;319-325

Amidation site:

Amino acids 229-233

287/550

FIGURE 287

CAGCGCTGACTGCGCCGCGGAGAAAGCCAGTGGGAACCCAGACCCATAGGAGACCCGCGTCCC
CGCTCGGCCTGGCCAGGCCCCGCGCTATGGAGTTCCTCTGGGCCCCCTCTCTTGGGTCTGTGCT
GCAGTCTGGCCGCTGCTGATCGCCACACCGTCTTCTGGAACAGTTCAAATCCCAAGTTCCGGA
ATGAGGACTACACCATACATGTGCAGCTGAATGACTACGTGGACATCATCTGTCCGCACTATG
AAGATCACTCTGTGGCAGACGCTGCCATGGAGCAGTACATACTGTACCTGGTGGAGCATGAGG
AGTACCAGCTGTGCCAGCCCCAGTCCAAGGACCAAGTCCGCTGGCAGTGCAACCGGCCAGTG
CCAAGCATGGCCCGGAGAAGCTGTCTGAGAAGTTCCAGCGCTTCACACCTTTACCCCTGGGCA
AGGAGTTCAAAGAAGGACACAGCTACTACTACATCTCCAAACCCATCCACCAGCATGAAGACC
GCTGCTTGAGGTTGAAGGTGACTGTCAGTGGCAAAATCACTCACAGTCCTCAGGCCCCATGACA
ATCCACAGGAGAAGAGACTTGACAGCAGATGACCCAGAGGTGCGGGTTCTACATAGCATCGGTC
ACAGTGCTGCCCCACGCCTCTTCCCACTTGCCCTGGACTGTGCTGCTCCTTCCACTTCTGCTGC
TGCAAACCCCCTGAAGGTGTGTGCCACACCTGGCCTTAAAGAGGGACAGGCTGAAGAGAGGGA
CAGGCACTCCAAACCTGTCTTGGGGCCACTTTCAGAGCCCCAGCCCTGGGAACCACTCCCAC
CACAGGCATAAGCTATCACCTAGCAGCCTCAAACCGGGTCAATATTAAGGTTTTCAACCGGAA
GGAGGCCAACCAGCCCGACAGTGCCATCCCCACCTTCACCTCGGAGGGATGGAGAAAGAAGTG
GAGACAGTCCTTTCCCACCATTCTGCCTTTAAGCCAAAGAAACAAGCTGTGCAGGCATGGTC
CCTTAAGGCACAGTGGGAGCTGAGCTGGAAGGGGCCACGTGGATGGGCAAAGCTTGTCAAAGA
TGCCCCCTTCAGGAGAGAGCCAGGATGCCCAGATGAACTGACTGAAGGAAAAGCAAGAAACAG
TTTCTTGCTTGGAAGCCAGGTACAGGAGAGGCAGCATGCTTGGGCTGACCCAGCATCTCCAG
CAAGACCTCATCTGTGGAGCTGCCACAGAGAAGTTTGTAGCCAGGTAAGTGCATTCTCTCCCAT
CCTGGGGCAGCACTCCCCAGAGCTGTGCCAGCAGGGGGGCTGTGCCAACCTGTTCTTAGAGTG
TAGCTGTAAGGGCAGTGCCCATGTGTACATTCTGCCTAGAGTGTAGCCTAAAGGGCAGGGCCC
ACGTGTATAGTATCTGTATATAAGTTGCTGTGTGTCTGTCTGATTCTACAAGTGGAGTTTT
TTTATACAATGTTCTTGTCTCAAATAAAGCAATGTGTTTTTCGG

288/550

FIGURE 288

MEFLWAPLLGLCCSLAAADRHTVFWNSSNPKFRNEDYTIHVQLNDYVDIICPHYEDHSADAAM
EQYILYLVEHEEYQLCQPQSKDQVRWQCNRPQSAKHGPEKLSEKFQRFPTFTLGKEFKEGHSYY
YISKPIHQHEDRCLRLKVTVSGKITHSPQAHDNPEKRLAADDPEVRVLHSIGHSAAPRLFPL
AWTVLLLPLLLLQTP

Important features:

Signal sequence:

Amino acids 1-17

N-glycosylation site:

Amino acids 26-30

Tyrosine kinase phosphorylation site:

Amino acids 118-127

N-myristoylation site:

Amino acids 10-16

289/550

FIGURE 289

CGGACGCGTGGGCGGACGCGTGGGCGGCCACGGCGCCCGGGCTGGGGCGGTGCGTTCTTC
CTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGGTAAAGATGGCCCCATGGCCCCGAAGGGC
CTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTTCCTCAACCTCCCAGGACCTATCTGGCTC
CAGCCCTCTCCACCTCCCCAGTCTTCTCCCCCGCCTCAGCCCCATCCGTGTCATACCTGCCGG
GGACTGGTTGACAGCTTTAACAAGGGCCTGGAGAGAACCATCCGGGACAACCTTTGGAGGTGGA
AACACTGCCTGGGAGGAAGAGAATTTGTCCAAATACAAAGACAGTGAGACCCGCCTGGTAGAG
GTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCTGAGTGAG
GAGCTGGTGGAGAGCTGGTGGTTTACAAGCAGCAGGAGGCCCCGGACCTCTTCCAGTGGCTG
TGCTCAGATTCCCTGAAGCTCTGCTGCCCCGCAGGCACCTTCGGGCCCTCCTGCCTTCCCTGT
CCTGGGGGAACAGAGAGGCCCTGCGGTGGCTACGGGCAGTGTGAAGGAGAAGGGACACGAGGG
GGCAGCGGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCTGTGGCCAGTGTGGCCTT
GGCTACTTTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTCCGGCTTGTTTTGGCCCCCTGT
GCCCCGATGCTCAGGACCTGAGGAATCAAACCTGTTTGCAATGCAAGAAGGGCTGGGCCCTGCAT
CACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAACTGTGGAGCTGACCAA
TTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAGGCCTGCCTAGGCTGC
ATGGGGGCAGGGCCAGGTCGCTGTAAGAAGTGTAGCCCTGGCTATCAGCAGGTGGGCTCCAAG
TGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGAACAAAGCAGTGTGAAAAC
ACCGAGGGCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAGATGGAAGGCATCTGTGTG
AAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGACAGAAGACGAGTTGGTGGTG
CTGCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCACGCTGGCTGCTAAGGGCGAC
TTGGTGTTACCGCCATCTTCATTGGGGCTGTGGCGGCCATGACTGGCTACTGGTTGTCAGAG
CGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGATAATCGCGGCCACCACCTGTAGGA
CCTCCTCCCACCCACGCTGCCCCAGAGCTTGGGCTGCCCTCCTGCTGGACACTCAGGACAGC
TTGGTTTATTTTGTAGAGTGGGGTAAGCACCCCTACCTGCCTTACAGAGCAGCCCAGGTACCC
AGGCCCCGGCAGACAAGGCCCTGGGGTAAAAAGTAGCCCTGAAGGTGGATACCATGAGCTCT
TCACCTGGCGGGGACTGGCAGGCTTCACAATGTGTGAATTTCAAAGTTTTCTTAATGGTG
GCTGCTAGAGCTTTGGCCCCCTGCTTAGGATTAGGTGGTCCTCACAGGGGTGGGGCCATCACAG
CTCCCTCCTGCCAGCTGCATGCTGCCAGTTCTGTGTTCTGTGTTACCACATCCCCACACCCCA
TTGCCACTTATTTATTCATCTCAGGAAATAAAGAAAGGTCTTGAAAGTTAAAAA
AAAAA

290/550

FIGURE 290

MAPWPPKGLVPAVLWGLSLFLNLP GPIWLQPSPPPQSSPPPQPHPCHTCRGLVDSFNKGLERT
IRDNFGGGNTAWEEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLLELSEELVESWWFHKQQE
APDLFQWLCSDSLKLCCPAGTFGPSCLP CGGTERPCGGYGQCEGEGTRGGSGHCDQCAGYGG
EACGQCGLGYFEARNASHLVCSACFGPCARCSGPEESNCLQCKKGWALHHLKCVDIDE CGTE
GANCGADQFCVNTEGSYECRDCAKACLGCMGAGPGRCKKCSPGYQQVGSKCLDVDECETEVC P
GENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTEDELVVLQQMFFGIIICAL
ATLAAKGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

Important features:**Signal sequence:**

Amino acids 1-29

Transmembrane domain:

Amino acids 342-392

N-glycosylation sites:

Amino acids 79-83;205-209

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 290-294

Aspartic acid and asparagine hydroxylation site:

Amino acids 321-333

EGF-like domain cysteine pattern signature:

Amino acids 181-193

291/550

FIGURE 291

CAGGTCCAACTGCACCTCGGTTCTATCGATTGAATCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGACCCAC
GCGTCCGAACACAGGTCCTTGTTGCTGCAGAGAAGCAGTTGTTTTGCTGGAAGGAGGGAGTGCGCGGGCTGCCCC
GGGCTCCTCCCTGCCGCCTCCTCTCAGTGGATGGTTCCAGGCACCTGTCTGGGGCAGGGAGGGCACAGGCCTGC
ACATCGAAGGTGGGGTGGGACCAGGCTGCCCTCGCCCCAGCATCCAAGTCTCCCTTGGGCGCCCGTGGCCCTG
CAGACTCTCAGGGCTAAGGTCCTCTGTTGCTTTTTGGTTCCACCTTAGAAGAGGCTCCGCTTGACTAAGAGTAGC
TTGAAGGAGGCACCATGCAGGAGCTGCATCTGCTCTGGTGGGCGCTTCTCCTGGGCCTGGCTCAGGCCTGCCCTG
AGCCCTGCGACTGTGGGGAAAAGTATGGCTTCCAGATCGCCGACTGTGCCTACCGCGACCTAGAATCCGTGCCGC
CTGGCTTCCCGGCCAATGTGACTACACTGAGCCTGTGAGCCAACCGGCTGCCAGGCTTGCCGGAGGGTGCCTTCA
GGGAGGTGCCCTGCTGCAGTCGCTGTGGCTGGCACACAATGAGATCCGCACGGTGGCCGCCGGAGCCCTGGCCT
CTCTGAGCCATCTCAAGAGCCTGGACCTCAGCCACAATCTCATCTCTGACTTTGCCTGGAGCGACCTGCACAACC
TCAGTGCCCTCCAATTGCTCAAGATGGACAGCAACGAGCTGACCTTCATCCCCCGGACGCCCTTCCGCAGCCTCC
GTGCTCTGCGCTCGCTGCAACTCAACCACAACCGCTTGACACATTGGCCGAGGGCACCTTCACCCCGCTCACCG
CGCTGTCCACCTGCAGATCAACGAGAACCCTTTCGACTGCACCTGCGGCATCGTGTGGCTCAAGACATGGGCCC
TGACCACGGCCGTGTCCATCCCGGAGCAGGACAACATCGCCTGCACCTCACCCCATGTGCTCAAGGGTACACCGC
TGAGCCGCCTGCCGCCACTGCCATGCTCGGCGCCCTCAGTGCAGCTCAGCTACCAACCCAGCCAGGATGGTGCCG
AGCTGCGGCCTGGTTTTGTGCTGGCACTGCACTGTGATGTGGACGGGCAGCCGGCCCTCAGCTTCACTGGCACA
TCCAGATACCCAGTGGCATTGTGGAGATCACCAGCCCCAACGTGGGCACTGATGGGCGTGCCCTGCCTGGCACCC
CTGTGGCCAGCTCCAGCCGCGCTTCCAGGCCTTTGCCAATGGCAGCCTGCTTATCCCGACTTTGGCAAGCTGG
AGGAAGGCACCTACAGCTGCCTGGCCACCAATGAGCTGGGCAGTGTGAGAGCTCAGTGGACGTGGCACTGGCCA
CGCCCGGTGAGGGTGGTGAGGACACACTGGGGCGCAGGTTCCATGGCAAAGCGGTTGAGGGAAAGGGCTGCTATA
CGGTTGACAACGAGGTGCAGCCATCAGGGCCGGAGGACAATGTGGTCATCATCTACCTCAGCCGTGCTGGGAACC
CTGAGGCTGCAGTCGCAGAAGGGGTCCCTGGGCAGCTGCCCCAGGCCTGCTCCTGCTGGGCCAAAGCCTCCTCC
TCTTCTTCTCCTCACCTCCTTCTTAGCCCCACCCAGGGCTTCCCTAACTCCTCCCCTTGCCCCTACCAATGCCCC
TTTAAGTGCTGCAGGGGTCTGGGGTTGGCAACTCCTGAGGCCTGCATGGGTGACTTCACATTTTCTACCTCTCC
TTCTAATCTCTTCTAGAGCACCTGCTATCCCCAATTCTAGACCTGCTCCAACTAGTGACTAGGATAGAATTTG
ATCCCCTAACTCACTGTCTGCGTGCTCATTGTCTGCTAACAGCATTGCCTGTGCTCTCCTCTCAGGGGCAGCATG
CTAACGGGGCGACGTCCTAATCCAATGGGAGAAGCCTCAGTGGTGGAATTCAGGCACTGTGACTGTCAAGCTG
GCAAGGGCCAGGATTGGGGGAATGGAGCTGGGGCTTAGCTGGGAGGTGGTCTGAAGCAGACAGGGAATGGGAGAG
GAGGATGGGAAGTAGACAGTGGCTGGTATGGCTCTGAGGCTCCCTGGGGCCTGCTCAAGCTCCTCCTGCTCCTTG
CTGTTTTCTGATGATTTGGGGCTTGGGAGTCCCTTTGTCTCATCTGAGACTGAAATGTGGGGATCCAGGATGG
CCTTCTTCTCCTTACCTTCTCCTCAGCCTGCAACCTCTATCCTGGAACCTGTCTCCTTTCTCCCCAACT
ATGCATCTGTTGTCTGCTCCTCTGCAAAGGCCAGCCAGCTTGGGAGCAGCAGAGAAATAAACAGCATTTCTGATG
CCAAAAAAAAAAAAAAAAAGGGCGCCGCGACTCTAGAGTCGACCT

292/550

FIGURE 292

MQELHLLWWALLLGLAQACPEPCDCGEKYGFQIADCAYRDLESVPPGFANVTTLSSLNRLP
GLPEGAFREVPLLQSLWLAHNEIRTVAAGALASLSHLKSLDLSHNLISDFAWSDLHNLSALQL
LKMSNELTFIPRDAFRSLRALRSLQLNHNRLHTLAEGTFTPLTALSHLQINENPFDCTCGIV
WLKTWALTAVSIPEQDNIACTSPHVLKGTPLSRLPPLPCSAPSVQLSYQPSQDGAELRPGFV
LALHCDVDGQPAPQLHWHIQIPSGIVEITSPNVGTDGRALPGTPVASSQPRFQAFANGSLLIP
DFGKLEEGTYSCLATNELGSAESSVDVALATPGEGEDTLGRRFHGKAVEGKGCYTVDNEVQP
SGPEDNVVIIYLSRAGNPEAAVAEGVPGQLPPGLLLLGQSLLLLFFFLTSF

Important features:**Signal peptide:**

amino acids 1-18

Transmembrane domain:

amino acids 403-418

N-glycosylation sites:

Amino acids 51-55, 120-124, 309-313

Tyrosine kinase phosphorylation site:

amino acids 319-326

N-myristoylation sites:amino acids 14-20, 64-70, 92-98, 218-224, 294-300, 323-329, 334-340,
350-356, 394-400**Amidation site:**

amino acids 355-359

Leucine Rich Repeat:

amino acids 51-74, 75-98, 99-122, 123-146, 147-170

Leucine rich repeat C-terminal domain:

amino acids 180-230

293/550

FIGURE 293

ACTTGGAGCAAGCGGCGGGCGGAGACAGAGGCAGAGGCAGAAGCTGGGGCTCCGTCTCGCCTCCCACGAGCG
ATCCCCGAGGAGAGCCGCGGCTCGGCGAGGCGAAGAGGCCGACGAGGAAGACCCGGGTGGCTGCGCCCTGCC
TCGCTTCCCAGGCGCGGCGGCTGCAGCCTTCCCCCTCTTGCTCGCCTTGAAAATGGAAAAAGATGCTCGCAGGCT
GCTTTCTGTGATCCTCGGACAGATCGTCTCTCCCTGCCGAGGCCAGGGAGCGGTACAGTGGGAGGTCCATCT
CTAGGGGCGAGACGCTCGGACCCACCCGAGACGGCCCTTCTGGAGAGTTCTGTGAGAACAAGCGGGCAGACC
TGGTTTTTCATCATTGACAGCTCTCGCAGTGTCAACACCCATGACTATGCAAAGGTCAAGGAGTTCATCGTGGACA
TCTTGCAATTCTTGACATTGGTCTGATGTCAACCGAGTGGGCTGCTCCAATATGGCAGCACTGTCAAGAATG
AGTTCTCCCTCAAGACCTTCAAGAGGAAGTCCGAGGTGGAGCGTGTGTCAAGAGGATGCGGCATCTGTCCACGG
GCACCATGACTGGGCTGGCCATCCAGTATGCCCTGAACATCGCATTCTCAGAAGCAGAGGGGGCCGCGCCCTGA
GGGAGAATGTGCCACGGGTGATAATGATCGTGACAGATGGGAGACCTCAGGACTCCGTGGCCGAGGTGGCTGCTA
AGGCACGGGACACGGGCATCCTAATCTTGCCATTGGTGTGGGCCAGGTAGACTTCAACACCTTGAAGTCCATTG
GGAGTGAGCCCCATGAGGACCATGTCTTCTGTGGCCAAATTCAGCCAGATTGAGACGCTGACCTCCGTGTTCC
AGAAGAAGTTGTGCACGGGCCACATGTGCAGCACCCCTGGAGCATAAAGTGTGCCCACTTCTGCATCAACATCCCTG
GCTCATACGTCTGCAGGTGCAACAAGGCTACATTCTCAACTCGGATCAGACGACTTGCAGAATCCAGGATCTGT
GTGCCATGGAGGACCACAACTGTGAGCAGCTCTGTGTGAATGTGCCGGGCTCCTTCGTCTGCCAGTGCTACAGTG
GCTACGCCCTGGCTGAGGATGGGAAGAGGTGTGTGGCTGTGGACTACTGTGCTCAGAAAACACCGGATGTGAAC
ATGAGTGTGTAATGCTGTGGCTCTACCTTTGCCAGTGCCATGAAGGATTTGCTCTTAACCCAGATGAAAAAA
CGTGACAAAGGATCAACTACTGTGCACTGAACAAACCGGGCTGTGAGCATGAGTGCGTCAACATGGAGGAGAGCT
ACTACTGCCGCTGCCACCGTGGCTACACTCTGGACCCCAATGGCAAAACCTGCAGCCGAGTGGACCACTGTGCAC
AGCAGGACCATGGCTGTGAGCAGCTGTGTCTGAACACGGAGGATTCTTCGTCTGCCAGTGCTCAGAAGGCTTCC
TCATCAACGAGGACCTCAAGACCTGCTCCCGGTGGATTACTGCCTGCTGAGTGACCATGGTTGTGAATACTCCT
GTGTCAACATGGACAGATCCTTTGCCCTGTCACTGTCTGAGGGACACGTGCTCCGACGATGGGAAGACGTGTG
CAAAATTGGACTCTTGCTCTGGGGGACCAAGGTTGTGAACATTCTGTGTGAAGCAGTGAAGATTCTGTTGTG
GCCAGTGCTTTGAAGTTATATACTCCGTGAAGATGGAAAAACCTGCAGAAGGAAGATGTCTGCCAAGCTATAG
ACCATGGCTGTGAACACATTTGTGTGAACAGTGACGACTCATACACGTGCGAGTGCTTGGAGGGATTCCGGCTCG
CTGAGGATGGGAAACGCTGCCGAAGGAAGGATGTCTGCAAAATCAACCCACCATGGCTGCCAACACATTTGTGTA
ATAATGGGAATTCCTACATCTGCAAAATGCTCAGAGGATTTGTTCTAGCTGAGGACGGAAGACGGTGCAAGAAAT
GCACTGAAGGCCCAATTGACCTGGTCTTTGTGATCGATGGATCCAAGAGTCTTGAGAGAAGAGAATTTTGAGGTCTG
TGAAGCAGTTTGTCACTGGAATTATAGATTCTTGACAATTTCCCCCAAAGCCGCTCGAGTGGGGCTGCTCCAGT
ATTCCACACAGGTCCACACAGAGTTCACTCTGAGAACTTCAACTCAGCCAAAGACATGAAAAAGCCGTGGCCC
ACATGAAATACATGGGAAAGGGCTCTATGACTGGGCTGGCCCTGAAACACATGTTTGAGAGAAGTTTACCACAAG
GAGAAGGGGGCCAGGCCCTTTCCACAAGGGTGCCAGAGCAGCCATTGTGTTCAACGACGACGGGCTCAGGATG
ACGTCTCCGAGTGGGCCAGTAAAGCCAAGGCCAATGGTATCACTATGTATGCTGTTGGGGTAGGAAAAGCCATTG
AGGAGGAACACAGAGATTGCCTCTGAGCCCAACAAACAGCATCTCTTCTATGCCGAAGACTTCAGCACAAATGG
ATGAGATAAGTGAAAAACTCAAGAAAGGCATCTGTGAAGCTCTAGAAGACTCCGATGGAAGACAGGACTCTCCAG
CAGGGGAACCTGCCAAAAACGGTCCAACAGCCAACAGAATCTGAGCCAGTCACCATAAATATCCAAGACCTACTTT
CCTGTTCTAATTTTGCAGTGCAACACAGATATCTGTTTGAAGAAGACAATCTTTTACGGTCTACACAAAAGCTTT
CCCATTCAACAAAACCTTCAGGAAGCCCTTTGGAAGAAAAACACGATCAATGCAAAATGTGAAAACCTTATAATGT
TCCAGAACCTTGCAACGAAGAAGTAAGAAAATTAACACAGCGCTTAGAAGAAATGACACAGAGAATGGAAGCCC
TGGAAATCGCTGAGATACAGATGAAGATTAGAAATCGCGACACATTTGTAGTCATTGTATCACGGATTACAAT
GAACGCACTGCAGAGCCCCAAGCTCAGGCTATTGTTAAATCAATAATGTTGTGAAGTAAACAATCAGTACTGA
GAAACCTGGTTTGCCACAGAACAAGACAAGAGTATACACTAATCTGTATAAATTTATCTAGGAAAAAATCCT
TCAGAATTCTAAGATGAATTTACCAGGTGAGAATGAATAAGCTATGCAAGGTATTTGTAATATACTGTGGACAC
AACTTGCTTCTGCCTCATCTGCCTTAGTGTGCAATCTCATTTGACTATACGATAAAGTTTGACAGTCTTACTT
CTGTAGAACACTGGCCATAGGAAATGCTGTTTTTTGTACTGGACTTTACCTTGATATATGTATATGGATGTATG
CATAAAATCATAGGACATATGTAATTGTGGAACAAGTTGGATTTTTTATACAATATTAATTCACCACTTCAG

294/550

FIGURE 294

MEKMLAGCFLILGQIVLLPAEARERSRGRSISRGRHARTHPTALLESSCENKRADLVFIID
SSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFRKRSEVERAVKR
MRHLSTGTMGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPDQDSVAEVAAKARDTGI
LIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAHMCSTLEHNCAHF
CINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYALAEDG
KRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCTRINYCALNKPGEHECVN
MEESYYCRCHRGYTLDPNGKTC SRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGF LINEDLKTC
SRVDYCLLSDHGCEYSCVNMDRSFACQCP EGHVLRSDGKTC AKLDSCALGDHGCEHSCVSSD
SFVCQCFEGYILREDGKTCRRKDVCAIDHGCEHICVNSDDSYTCECLEGFRLAEDGKRCRRK
DVCKSTHHGCEHICVNNNGNSYICKCEGFVLAEDGRRCKKCTEGPIDLVFVIDGSKSLGEENF
EVVKQFVTGLIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNSAKDMKKAVAHMKYMGKGSMT
GLALKHMFERSFTQGEGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKANGITMYAVGVGKA
IEEELQEIASEPTNKHLYFAEDFSTMDEISEKLKKGICEALEDSDGRQDSPAGELPKTVQQPT
ESEPVTINIQDLLSCSNFAVQHRYLFEEENLLRSTQKLSHSTKPSGSPLEEKHDQCKCENLIM
FQNLANEEVRKLTQRLEEMTQRM EALENRLRYR

Important features:**Signal sequence:**

Amino acids 1-23

N-glycosylation site:

Amino acids 221-225

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 115-119; 606-610; 892-896

N-myristoylation sites:Amino acids 133-139; 258-264; 299-305; 340-346; 453-459; 494-500;
639-645; 690-694;
752-758; 792-798**Amidation sites:**

Amino acids 314-318; 560-564; 601-605

Aspartic acid and asparagine hydroxylation sites:Amino acids 253-265; 294-306; 335-347; 376-388; 417-429;
458-470; 540-552; 581-593

295/550

FIGURE 295

GGCCGGAGCAGCACGGCCGCAGGACCTGGAGCTCCGGCTGCGTCTTCCCGCAGCGCTACCCGC
C**ATG**CGCCTGCCGCGCCGGGCCGCGCTGGGGCTCCTGCCGCTTCTGCTGCTGCTGCCGCCCCG
GCCGGAGGCCGCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGCTGGTGGACAAGTTTAA
CCAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGGCGGGAACACGGCTTGGGAGGAAAA
GACGCTGTCCAAGTACGAGTCCAGCGAGATTGCGCTGCTGGAGATCCTGGAGGGGCTGTGCGA
GAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGAGCACCTGGAGGCCTGGTG
GCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTTGTGTGAAGACACTGAAAGT
GTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGCCAGGGCGGATCCCAGAGGCC
CTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACAGGGCGACGGGTCTGCCGGTG
CCACATGGGGTACCAGGGCCCCGCTGTGCACTGACTGCATGGACGGCTACTTCAGCTCGCTCCG
GAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCCTGCAAGACGTGCTCGGGCCTGAC
CAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCTGGACGAGGGCGCCTGTGTGGATGT
GGACGAGTGTGCGGCCGAGCCGCCTCCCTGCAGCGCTGCGCAGTTCTGTAAGAACGCCAACGG
CTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGTGGGCTGCACAGGGGAAGGCCCAGGAAA
CTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCACGGACAGTGTGCAGATGTGGACGAGTG
CTCACTAGCAGAAAAAACCTGTGTGAGGAAAAACGAAACTGCTACAATACTCCAGGGAGCTA
CGTCTGTGTGTCTGACGGCTTCGAAGAAACGGAAGATGCCTGTGTGCCGCCGGCAGAGGC
TGAAGCCACAGAAGGAGAAAGCCCGACACAGCTGCCCTCCCGCGAAGACCTG**TAA**TGTGCCGT
ACTTACCCTTTAAATTATTCAGAAGGATGTCCCGTGGAAAATGTGGCCCTGAGGATGCCGTCT
CCTGCAGTGGACAGCGGCGGGGAGAGGCTGCCTGCTCTTAACGGTTGATTCTCATTTGTCCC
TTAAACAGCTGCATTTCTTGTTGTTCTTAAACAGACTTGTATATTTTGATACAGTTCTTTGT
AATAAAATTGACCATTGTAGGTAATCAGGAGGAAAAA

296/550

FIGURE 296

MRLPRRAALGLLPLLLLLPPAPEAAKKPTPCHRCRGLVDKFNQGMVDTAKKNFEGGGNTAWEEK
TLSKYESSEIRLLEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLFEWFCVKTLKV
CCSPGTYGPDCCLACQGGGSRPCSGNGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLR
NETHSICTACDESKTCSGLTNRDCGECEVGVWLDEGACVDVDECAAEPFPCSAAQFCCKNANG
SYTCEECDSSCVGCTGEGPGNCKECISGYAREHGQCADVDECSLAEKTCVRKNENCYNTPGSY
VCVCPDGFEEETEDACVPPAEAEATEGESPTQLPSREDL

Important features:**Signal peptide:**

Amino acids 1-24

N-glycosylation sites:

Amino acids 190-194;251-255

Glycosaminoglycan attachment sites:

Amino acids 149-153;155-159

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 26-30

Tyrosine kinase phosphorylation site:

Amino acids 303-310

N-myristoylation sites:Amino acids 44-50;54-60;55-61;81-87;150-156;158-164;164-170;
252-258;313-319**Aspartic acid and asparagine hydroxylation site:**

Amino acids 308-320

EGF-like domain cysteine pattern signature:

Amino acids 166-178

Leucine zipper pattern:

Amino acids 94-116

297/550

FIGURE 297

GACATCGGAGGTGGGCTAGCACTGAACTGCTTTTCAAGACGAGGAAGAGGAGGAGAAAGAGAAAGAAGAGGAAG
ATGTTGGGCAACATTTATTTAACATGCTCCACAGCCCGGACCCTGGCATCATGCTGCTATTCCTGCAAATACTGA
AGAAGCATGGGATTTAAATATTTTACTTCTAAATAAATGAATTACTCAATCTCCTATGACCATCTATACATACTC
CACCTTCAAAAAGTACATCAATATTATATCATTAAAGGAAATAGTAACCTTCTCTTCTCCAATATGCATGACATTT
TTGGACAATGCAATTGTGGCACTGGCACTTATTTTCAGTGAAGAAAACTTTGTGGTTCTATGGCATTTCATCATTT
GACAAATGCAAGCATCTTCCTTATCAATCAGCTCCTATTGAACCTACTAGCACTGACTGTGGAATCCTTAAGGGC
CCATTACATTTCTGAAGAAGAAAGCTAAGATGAAGGACATGCCACTCCGAATTCATGTGCTACTTGGCCTAGCTA
TCACTACACTAGTACAAGCTGTAGATAAAAAAGTGGATTGTCCACGGTTATGTACGTGTGAAATCAGGCCTTGGT
TTACACCCAGATCCATTTATATGGAAGCATCTACAGTGGATTGTAATGATTTAGGTCTTTTAACTTTCCAGCCA
GATTGCCAGCTAACACACAGATTCTTCTCCTACAGACTAACAAATATTGCAAAAATTGAATACTCCACAGACTTTC
CAGTAAACCTTACTGGCCTGGATTTATCTCAAAACAATTTATCTTCAGTCACCAATATTAATGTAAAAAAGATGC
CTCAGCTCCTTTCTGTGTACCTAGAGGAAAACAACTTACTGAACCTGCCTGAAAAATGTCTGTCCGAACCTGAGCA
ACTTACAAGAACTCTATATTAATCACAACCTTGCTTTCTACAATTTACCTGGAGCCTTTATTGGCCTACATAATC
TTCTTCGACTTCATCTCAATTCAAATAGATTGCAGATGATCAACAGTAAGTGGTTTGATGCTCTTCCAAATCTAG
AGATTCTGATGATTGGGGAAAATCCAATTATCAGAATCAAAGACATGAACCTTTAAGCCTCTTATCAATCTTCGCA
GCCTGGTTATAGCTGGTATAAACCTCACAGAAATACCAGATAACGCCTTGGTTGGACTGGAAAACCTTAGAAAGCA
TCTCTTTTACGATAACAGGCTTATTAAAGTACCCCATGTTGCTCTTCAAAAAGTTGTAAATCTCAAATTTTTGG
ATCTAAATAAAAATCCTATTAATAGAATACGAAGGGGTGATTTAGCAATATGCTACACTTAAAAGAGTTGGGGA
TAAATAATATGCCTGAGCTGATTCCATCGATAGTCTTGCTGTGGATAACCTGCCAGATTTAAGAAAAATAGAAG
CTACTAACAACCCTAGATTGTCTTACATTCACCCCAATGCATTTTTTCAGACTCCCCAAGCTGGAATCACTCATGC
TGAACAGCAATGCTCTCAGTGCCTGTACCATGGTACCATTGAGTCTCTGCCAAACCTCAAGGAAATCAGCATAAC
ACAGTAACCCCATCAGGTGTGACTGTGTATCCGTTGGATGAACATGAACAAAACCAACATTTCGATTTCATGGAGC
CAGATTCACTGTTTTGCGTGGACCCACCTGAATTCGAAGTTCAGAAATGTTCCGCAAGTGCATTTTCAGGGACATGA
TGGAATTTGTCTCCCTCTTATAGCTCCTGAGAGCTTTCTCTAATCTAAATGTAGAAGCTGGGAGCTATGTTT
CCTTTCACTGTAGAGCTACTGCAGAACCACAGCCTGAAATCTACTGGATAACACCTTCTGGTCAAAAACCTCTTGC
CTAATACCCTGACAGACAAGTTCTATGTCCATTCTGAGGGAACACTAGATATAAATGGCGTAACTCCCAAGAAG
GGGTTTATATACTTGTATAGCAACTAACCTAGTTGGCGCTGACTTGAAGTCTGTTATGATCAAAGTGGATGGAT
CTTTTCCACAAGATAACAATGGCTCTTTGAATATTAATAAAGAGATATTCAGGCCAATTCAGTTTGGTGTCTT
GGAAAGCAAGTTCTAAAATCTCAAATCTAGTGTTAAATGGACAGCCTTTGTCAAGACTGAAAATCTCATGCTG
CGCAAAGTGCTCGAATACCATCTGATGTCAAGGTATATAATCTTACTCATCTGAATCCATCAACTGAGTATAAAA
TTTGTATTGATATTTCCACCCTATCAGAAAAACAGAAAAAATGTGTAAATGTCACCACCAAGGTTTGCACC
CTGATCAAAAAGAGTATGAAAAGAATAATACCACAACACTTATGGCCTGTCTTGGAGGCCTTCTGGGGATTATTG
GTGTGATATGTCTTATCAGCTGCCTCTCTCCAGAAATGAACCTGTGATGGTGGACACAGCTATGTGAGGAATTACT
TACAGAAACCAACCTTTGCATTAGGTGAGCTTTATCCTCCTCTGATAAATCTCTGGGAAGCAGGAAAAAGAAAAA
GTACATCACTGAAAGTAAAAGCAACTGTTATAGGTTTACCAACAAATATGTCCTAAAAAACCACCAAGGAAACCTA
CTCCAAAATGAAC

298/550

FIGURE 298

MKDMPLRIHVLLGLAITTLVQAVDKKVDPCRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLTF
PARLPANTQIILLQTNNAKIEYSTDFPVNLTGLDLSQNNLSSVTNINVKKMPQLLSVYLEEN
KLTPEKCLSELNLQELYINHNLSTISPGAFIHLNLLRLHLNSNRLQMINSKWFDALPN
LEILMIGENPIIRIKDMNFKPLINLRSLVIAGINLTEIPDNALVGLENLESISFYDNRLIKVP
HVALQKVVLNKFLLDNKNPINRIRRGDFSNNMLHLKELGINNMPELISIDSLAVDNLPDLRKIE
ATNNPRLSYIHPNAFFRLPKLESMLNSNALSALYHGTTIESLPNLKEISIHSPNIRCDVCVIRW
MNMNKTNIRFMEPDSLFCVDPPEFQGGQNVVRQVHFRDMMEICLPLIAPESFPSNLNVEAGSYVS
FHCRTAEAPQPEIYWITPSGQKLLPNTLTDFYVHSEGTLDINGVTPKEGGLYTCIATNLVGA
DLKSVMIKVDGSFPQDNNGSLNIKIRDIQANSVLVSWKASSKILKSSVKWTAFVKTENSHAAQ
SARIPSDVKVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVTTKGLHPDQKEYEKNNTTTLM
ACLGGLLGIIGVICLISCLSPENMCDGGHSYVRNYLQKPTFALGELYPPLINLWEAGKEKSTS
LKVKATVIGLPTNMS

Important features:**Signal sequence:**

amino acids 1-22

Transmembrane domain:

amino acids 633-650

N-glycosylation site.amino acids 93-97, 103-107, 223-227, 382-386, 522-526, 579-583,
608-612, 624-628, 625-629**Casein kinase II phosphorylation site.**

amino acids 51-55, 95-99, 242-246, 468-472, 487-491

Tyrosine kinase phosphorylation site.

amino acids 570-579

N-myristoylation site.amino acids 13-19, 96-102, 158-164, 221-227, 352-358, 437-443,
491-497, 492-498, 634-640, 702-708**Cell attachment sequence.**

amino acids 277-280

299/550

FIGURE 299

GCTGTGGGAACCTCTCCACGCGCACGAACTCAGCCAACGATTTCTGATAGATTTTTGGGAGTT
TGACCAGAGATGCAAGGGGTGAAGGAGCGCTTCTACCGTTAGGGAACTCTGGGGACAGAGCG
CCCCGGCCGCCTGATGGCCGAGGCAGGGTGCGACCCAGGACCCAGGACGGCGTCGGGAACCAT
ACCATGGCCCCGGATCCCCAAGACCCTAAAGTTTCGTCTGTCATCGTCGCGGTCTGCTGCCA
GTCCTAGCTTACTCTGCCACCACTGCCCCGGCAGGAGGAAGTTCCCCAGCAGACAGTGGCCCCA
CAGCAACAGAGGCACAGCTTCAAGGGGGAGGAGTGTCCAGCAGGATCTCATAGATCAGAACAT
ACTGGAGCCTGTAACCCGTGCACAGAGGGTGTGGATTACACCAACGCTTCCAACAATGAACCT
TCTTGCTTCCCATGTACAGTTTGTAAATCAGATCAAAAACATAAAAGTTTCCTGCACCATGACC
AGAGACACAGTGTGTCAAGTGTAAAGAAGGCACCTTCCGGAATGAAACTCCCCAGAGATGTGC
CGGAAGTGTAGCAGGTGCCCTAGTGGGGAAGTCCAAGTCAGTAATTGTACGTCCTGGGATGAT
ATCCAGTGTGTTGAAGAATTTGGTGCCAATGCCACTGTGGAAACCCAGCTGCTGAAGAGACA
ATGAACACCAGCCCGGGGACTCCTGCCCCAGCTGCTGAAGAGACAATGAACACCAGCCCAGGG
ACTCCTGCCCCAGCTGCTGAAGAGACAATGACCACCAGCCCCGGGACTCCTGCCCCAGCTGCT
GAAGAGACAATGACCACCAGCCCGGGGACTCCTGCCCCAGCTGCTGAAGAGACAATGACCACC
AGCCCGGGGACTCCTGCCTCTTCTCATTACCTCTCATGCACCATCGTAGGGATCATAGTTCTA
ATTGTGCTTCTGATTGTGTTTGTTGAAAGACTTCACTGTGGAAGAAATTCCTTCCTTACCTG
AAAGGTTTCAGGTAGGCGCTGGCTGAGGGCGGGGGCGCTGGACACTCTCTGCCCTGCCTCCCT
CTGCTGTGTTCCACAGACAGAAACGCCTGC

300/550

FIGURE 300

MARI PKTLKFVVVIVAVLLPV LAYSATTARQEEVPQQT VAPQQQRHSFKGEECPAGSHRSEHT
GACNPCTEGVDYTNASNNEPSCFPCTVCKSDQKHKSSCTMTRDTVCQCKEGTFRNENSPEMCR
KCSRCPSGEVQVSNCTSWDDIQCVEEFGANATVETPAAEETMNTSPGTPAPAAEETMNTSPGT
PAPAAEETMTTSPGTPAPAAEETMTTSPGTPAPAAEETMTTSPGTPASSHYLSCTIVGIIIVLI
VLLIVFV

Important features:**Signal peptide:**

Amino acids 1-29

Transmembrane domain:

Amino acids 240-259

N-glycosylation site:

Amino acids 77-81;140-144;156-160

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 126-130

N-myristoylation sites:

Amino acids 56-62;72-78;114-120;154-160;233-239

301/550

FIGURE 301

CACAAGCATCTTAATTTGAATCCACAAAGTTTCATGTAATGAAAAGAAATACATAATTTTAAT
TCAACCCGAGTGTTTTCCAAGAAGATTGTATTTGCTTAAATTGCTACAGTAATTCAAGAGACA
GCCCTGTCTGGACACAGAGTTACTGTGGATTTTTAAGAGACTCAGTTAAAGAATTTAGGAATT
TCTGATTCATTTAAAGGATTTACAAATTCATCAACCCCTGAAAATAAGCAAATTGAACAGG
AAAAAAAAAAGAAGATGGGTTTTTTAAGTCCAATATATGTTATTTTCTTCTTTTTTGGAGTC
AAAGTACATTGCCAATATGAACTTATCAGTGGGATGAAGACTATGACCAAGAGCCAGATGAT
GATTACCAAACAGGATTCCCATTTTCGTCAAAATGTAGACTACGGAGTTCCTTTTCATCAGTAT
ACTTTAGGCTGTGTGCTGAGTGAATGCTTCTGTCCAATACTTTCCATCATCAATGTACTGTGAT
AATCGCAAATCAAGACTATCCCAAATATTCGATGCACATTCAGCAACTCTACCTTCAGTTC
AATGAAATTGAGGCTGTGACTGCAAATTCATTCATCAATGCAACTCATCTAAAGAAATTAAC
CTCAGCCACAACAAAATTAAATCTCAAAGATTGATTATGGTGTGTTTGCTAAGCTTCCAAAT
CTACTACAATTTCATCTAGAGCATAATAATTTAGAAGAATTTCCATTTCTCTTCTCTAAATCT
CTGGAAAGACTCCTTCTTGTTTACAATGAAATCTCCAACTGCAGACAAATGCTATGGATGGG
CTAGTAACTTGACCATGCTTGATCTCTGTTATAATTATCTTCATGATTCTCTGCTAAAAGAC
AAAATCTTTGCCAAAATGGAAAACTAATGCAGCTCAACCTCTGCAGTAACAGATTAGAATCA
ATGCCTCCTGGTTTGCTTCTTCACTTATGTATCTGTCTTTAGAAAATAATTCAATTTCTTCT
ATACCCGAAAAATACTTCGACAACTTCCAAAATTCATACTCTAAGAATGTACACAACAAA
CTACAAGACATCCCATATAATATTTTTAATCTTCCCAACATTGTAGAACTCAGTGTGGACAC
AACAAATTGAAGCAAGCATTCTATATTCCAAGAAATTTGGAACACCTATACCTACAAAATAAT
GAAATAGAAAAGATGAATCTTACAGTGATGTGTCCTTCTATTGACCCACTACATTACCACCAT
TTAACATACATTCGTGTGGACCAAAATAAACTAAAAGAACCAATAAGCTCATAACATCTTCTTC
TGCTTCCCTCATATACACACTATTTATTATGGTGAACAACGAAGCACTAATGGTCAAACAATA
CAACTAAAGACACAAGTTTTTCAGGAGATTTCCAGATGATGATGATGAAAGTGAAGATCACGAT
GATCCTGACAATGCTCATGAGAGCCCAGAACAAGAAGGAGCAGAAGGGCACTTTGACCTTCAT
TATTATGAAAATCAAGAATAGCAAGAACTATATAGGTATACACTTACGACTTCACAAAACCTA
TACTTAATATAGTAAATCTAAGTAAACATGTATTACTCAAAGTAATATATTTAGAATTATGTA
TTAGTATAAGATCAGAATTGAATTTAAGTTGTTGGTGACATCTGCATCATTTTCATAGGATTAG
AACTTACTCAAATAATGTAAATCTTTAAAAATATAAATTAGAATGACAAGTGGGAATCATAA
ATTAAACGTTAATGGTTTCTTATGCTCTTTTTAAATATAGAAATATCATGTAAAGAAAAAA
AAAAAA

302/550

FIGURE 302

MGFLSPIYVIFFFFGVKVHCQYETYQWDEDYDQEPDDDYQTGFPPFRQNVGYGVPFHQYTLGCV
SECFCPTNFPSSMYCDNRKLKTIPNIPMHIQQLYLQFNEIEAVTANSFINATHLKEINLSHNK
IKSQKIDYGVFAKLPNLLQLHLEHNNLEEFPPPLPKSLERLLLGYNEISKLTNAMDGLVNLT
MLDLCYNYLHDSLLKDKIFAKMEKLMQLNLCSNRLESMPPGLPSSLMYLSLENNSSISSEPEKY
FDKLPKLHTLRMSHNKLQDIPYNIFNLPNIVELSVGHNKLKQAFYIPRNLEHLYLQNNIEIEKM
NLTVMCPSIDPLHYHHLTYIRVDQNKLEPISSYIFFCFPHIHTIYYGEQRSTNGQTIQLKTQ
VFRFPDDDDDESEDHDDPDNAHESPEQEGAEGHFDLHYENQE

Important features:**N-glycosylation sites:**

Amino acids 113-117;121-125; 187-191;242-246;316-320

Tyrosine kinase phosphorylation sites:

Amino acids 268-275;300-307

N-myristoylation site:

Amino acids 230-236

Leucine zipper patterns:

Amino acids 146-168;217-239

303/550

FIGURE 303

CCCCGGGACTGGCGCAAGGTGCCCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTTAGC
TGCAGCCTTTTGAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTACCA
CGCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGCATGAATCTGGT
AGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTTGTTCCTTATGAT
ACTGTGCTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTTCCTCTGGGGGTTT
AAATGTCACCTGTAGCAATGCAATCTCAAGGAAATACCTAGAGATCTTCCTCCTGAAACAGT
CTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTTAAGGACCTCCA
TCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCATGCCTTCAA
AGGAGTAGCTGAAACCTTGCAACTCTGGACTTGTCCGACAATCGGATTCAAAGTGTGCACAA
AAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACAACCCCTGGCACTGCGACTG
TACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCCACAACGTGATCTG
TAAAACGTCCGTGTTGGATGAACATGCTGGCAGACCATTCCCTCAATGCTGCCAACGACGCTGA
CCTTTGTAACCTCCCTAAAAAACTACCGATTATGCCATGCTGGTCACCATGTTTGGCTGGTT
CACTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAAATCAGGAGGATGCCCCGAGACA
CCTCGAATACTTGAAATCCCTGCCAAGCAGGCAGAAGAAAGCAGATGAACCTGATGATATTAG
CACTGTGGTATAGTGTCCAACTGACTGTCATTGAGAAAGAAAGAAAGTAGTTTGCGATTGCA
GTAGAAATAAGTGGTTTACTTCTCCCATCCATTGTAAACATTTGAACTTTGTATTTTCAGTTT
TTTTTGAATTATGCCACTGCTGAACTTTTAACAAACACTACAACATAAATAATTTGAGTTTAG
GTGATCCACCCCTTAATTGTACCCCGATGGTATATTTCTGAGTAAGCTACTATCTGAACATT
AGTTAGATCCATCTCACTATTTAATAATGAAATTTATTTTTTTAATTTAAAAGCAAATAAAAG
CTTAACCTTTGAACCATGGGAAAAAAAAAAAAAAAAAAAAAACA

304/550

FIGURE 304

MNLVDLWLTRSLSMCLLLQS FVLMILCFHSASMCPKGCLCSSSGGLNVTCSNANLKEIPRDLP
PETVLLYLDNSNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGVAETLQTLDSLSDNRIQ
SVHKNAFNNLKARARIANNPWHC DCTLQQVLRSMASNHETAHNVICKTSVLDEHAGRPFLNAA
NDADLCNLPPKKT TDYAMLVTMFGWFTMVISYVVYYVRQNQEDARRHLEYLKS LPSRQKKADEP
DDISTVV

Important features:**Signal sequence:**

Amino acids 1-33

Transmembrane domain:

Amino acids 204-219

N-glycosylation sites:

Amino acids 47-51;94-98

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 199-203

Casein kinase II phosphorylation site.

amino acids 162-166, 175-179

N-myristoylation sites:

Amino acids 37-43;45-51;110-116

305/550

FIGURE 305

CGCCACCCTGCGGCCACCGCCAATGAAACGCCTCCCGCTCCTAGTGGTTTTTCCACTTTGTTGAATTGTTCCCT
ATACTCAAAATTGCACCAAGACACCTTGCTCTCCCAAATGCAAAATGTGAAATACGCAATGGAATTGAAGCCTGCT
ATTGCAACATGGGATTTTCAGGAAATGGTGTACAAATTTGTGAAGATGATAATGAATGTGGAAATTTAACTCAGT
CCTGTGGCGAAAATGCTAATTGCACTAACACAGAAGGAAGTTATTATTGTATGTGTACCTGGCTTCAGATCCA
GCAGTAACCAAGACAGGTTTATCACTAATGATGGAACCGTCTGTATAGAAAATGTGAATGCAAACTGCCATTTAG
ATAATGTCTGTATAGCTGCAAAATATTAATAAACTTTAACAAAAATCAGATCCATAAAAGAACCTGTGGCTTTGC
TACAAGAAGTCTATAGAAATCTGTGACAGATCTTACCAACAGATATAATTACATATATAGAAATATTAGCTG
AATCATCTTCTACTAGGTTACAAGAACAACACTATCTCAGCCAAGGACACCCTTCTAACTCAACTCTTACTG
AATTTGTA AAAACCGTGAATAATTTTGTTC AAGGGATACATTTGTAGTTTGGGACAAGTTATCTGTGAATCATA
GGAGAACACATCTTACAAAACCTCATGCACACTGTTGAACAAGCTACTTTAAGGATATCCAGAGCTTCCAAAAGA
CCACAGAGTTTGATACAAATTCACCGGATATAGCTCTCAAAGTTTCTTTTTTGTTCATATAACATGAAACATA
TTCATCCTCATATGAATAGGTGAGACTACATAAATATATTTCCAAAGAGAAAAGCTGCATATGATTCAAATG
GCAATGTTGCAGTTGCATTTTATATTATAAGAGTATTGGTCCCTTGGCTTTCATCATCTGACAACCTCTTATTGA
AACCTCAAAATATGATAATTTCTGAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTCAGTCTCAATGAGCTCAA
ACCCACCCACATTATATGAACCTTGAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATAGGTATAGGA
GTCTATGTGCATTTTGGAAATTA CTACCTGATACCATGAATGGCAGCTGGTCTTCAGAGGGCTGTGAGCTGACAT
ACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCTGACACATTTTGAATTTTGATGTCCTCTGGTCTCTT
CCATTGGTATTAAAGATTATAATATTCTTACAAGGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTG
CCATATGCATTTTACCTTCTGGTCTTTCAGTGAAATTC AAGCACCAGGACAACAATTCACAAAAATCTTTGCT
GTAGCCTATTTCTTGCTGAACCTGTTTTTCTGTTGGGATCAATACAAATACTAATAAGCTCTTCTGTTCAATCA
TTGCCGGACTGCTACACTACTTCTTTTTAGCTGCTTTTGCATGGATGTGCATTGAAGGCATACATCTCTATCTCA
TTGTTGTGGGTGTCATCTACAACAAGGGATTTTGCACAAGAATTTTATATCTTTGGCTATCTAAGCCCAGCCG
TGGTAGTTGGATTTTCGGCAGCACTAGGATACAGATATTATGGCACAACCAAGTATGTTGGCTTAGCACCCGAAA
ACAACTTTATTTGGAGTTTATAGGACCAGCATGCCTAATCATTCTTGTAAATCTCTTGGCTTTTGGAGTCATCA
TATACAAAGTTTTTCGTCACACTGCAGGGTTGAAACCAGAGTTAGTTGCTTTGAGAACATAAGGTCTTGTGCAA
GAGGAGCCCTCGCTCTTCTGTTCTCTCGGCACCACCTGGATCTTTGGGGTCTCCATGTTGTGCACGCATCAG
TGGTTACAGCTTACCTCTTACAGTCAGCAATGCTTTCCAGGGGATGTTCAATTTTTTATTCCTGTGTGTTTTAT
CTAGAAAGATTCAAGAAGAATATTACAGATTGTTCAAAAATGTCCCCTGTGTTTTGGATGTTTAAAGGTAAACAT
AGAGAATGGTGGATAATTACAACCTGCACAAAAATAAAAATCCAAGCTGTGGATGACCAATGTATAAAAATGACT
CATCAAATTATCCAATTATTAATACTAGACAAAAAGTATTTTAAATCAGTTTTTCTGTTTATGCTATAGGAACT
GTAGATAATAAGGTAAAATATGTATCATATAGATATACTATGTTTTCTATGTGAATAGTTCTGTCAAAAATA
GTATTGCAGATATTTGGAAAGTAATTGGTTTCTCAGGAGTGATATCACTGCACCCAAGGAAAGATTTTCTTTCTA
ACACGAGAAGTATATGAATGTCCTGAAGGAAACCACTGGCTTGATATTTCTGTGACTCGTGTGCTTTGAAACT
AGTCCCTTACCACCTCGGTAATGAGCTCCATTACAGAAAGTGGAACATAAGAGAATGAAGGGGCAGAATATCAAA
CAGTGAAAAGGGAATGATAAGATGTATTTGAATGAACTGTTTTTCTGTAGACTAGCTGAGAAATTTGTGACAT
AAAATAAAGAATTGAAGAAACACATTTTACCATTTTGTGAATTTGTTCTGAACTTAAATGTCCACTAAAACAACCTT
AGACTTCTGTTTGCTAAATCTGTTTCTTTTCTAATATTCTAAAAAAAAGGTTTACCTCCACAAATTGA
AA

306/550

FIGURE 306

MKRLPLLVVSTLLNCSYTNCTKTPCLPNAKCEIRNGIEACYCNMGFSNGVGTICEDDNECGNLTQSCGENANC
TNTEGSYYCMCVPGFRSSSNQDRFITNDGTVCIENVNANCHLDNVCIANINKTLTKIRSIKEPVALLQEVYRNS
VTDLSPTDIITYIEILAESSLLGYKNNTISAKDTLSNSTLTFVKTVNNFVQRDTFVVWDKLSVNHRRTHLTKL
MHTVEQATLRISQSFSQKTTEFDNSTDIALKVFFFDSDYNMKHIHPHMNDGDYINIFPKRKAAYDSNGNVAVAF
YYKSIGPLLSSSDNFLKPKQNYDNSEEEERVISVISVSMSSNPPTLYELEKITFTLSHRKVTDYRSLCAFWNY
SPDTMNGSWSSEGCELTYSNETHTSCRCNHLTHFAILMSSGSPSIGIKDYNILTRITQLGIIISLICLAICIFTFW
FFSEIQSTRTTIHKNLCCSLFLAELVFLVGINTNTNKLFCSSIIAGLLHYFFLAFAWMCIIEGHLVYLVVGVYIN
KGFLHKNFYIFGYLSPAVVVGFSALGYRYYGTTKVCWLSTENNFIWSFIGPACLIILVNLLAFGVIIYKVFRHT
AGLKPEVSCFENIRSCARGALALLFLGTTWIFGVLVHVVHASVVTAYLFTVSNAFQGMFIFLFLCVLSRKIQEEY
YRLFKNVPCCFGCLR

Important features:**Signal peptide:**

Amino acids 1-19

Transmembrane domain:

Amino acids 431-450;494-515;573-594;619-636;646-664

N-glycosylation sites:Amino acids 15-19;21-25;64-68;74-78;127-131;177-181;
188-192;249-253;381-385;395-399**Glycosaminoglycan attachment site:**

Amino acids 49-53

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 360-364

Tyrosine kinase phosphorylation sites:

Amino acids 36-44;670-677

N-myristoylation sites:Amino acids 38-44;50-56;52-58;80-86;382-388;388-394;
434-440;480-486;521-527**Aspartic acid and asparagine hydroxylation site:**

Amino acids 75-87

307/550

FIGURE 307

CCAGGCCGGGAGGCGACGCGCCAGCCGTCTAAACGGGAACAGCCCTGGCTGAGGGAGCTGCAGCGCAGCAGAGT
ATCTGACGGCGCCAGGTTGCGTAGGTGCGGCACGAGGAGTTTTCCCGGCAGCGAGGAGGTCCTGAGCAGC**CATGGC**
CCGGAGGAGCGCCTTCCCTGCCGCGCGCTCTGGCTCTGGAGCATCCTCCTGTGCCTGCTGGCACTGCGGGCGGA
GGCCGGGCGCCGAGGAGGAGCCTGTACCTATGGATCGATGCTACCAGGCAAGAGTACTCATAGGATTTGA
AGAAGATATCCTGATTGTTTCAGAGGGGAAAATGGCACCTTTTACACATGATTTTCAGAAAAGCGCAACAGAGAAT
GCCAGCTATTCCTGTCAATATCCATTCCATGAATTTTACCTGGCAAGCTGCAGGGCAGGCAGAATACTTCTATGA
ATTCTGTCTCTGCGCTCCCTGGATAAAGGCATCATGGCAGATCCAACCGTCAATGTCCCTCTGCTGGGAACAGT
GCCTCACAAGGCATCAGTTGTTCAAGTTGGTTTCCCATGTCTTGGAACACAGGATGGGGTGGCAGCATTTGAAGT
GGATGTGATTGTTATGAATTTGAAGGCAACACCATTTCTCAAACACCTCAAATGCTATCTTCTTTAAACATG
TCAACAAGCTGAGTGCCCGAGGCGGGTGCCGAAATGGAGGCTTTTGTAAAGACGCATCTGCGAGTGTCTCTGA
TGGGTTCCACGGACCTCACTGTGAGAAAGCCCTTTGTACCCACGATGTATGAATGGTGGACTTTGTGTGACTCC
TGGTTTTCTGCATCTGCCCACCTGGATTCTATGGAGTGAAGTGTGACAAAGCAAAGTGTCAACCACTGCTTTAA
TGGAGGGACCTGTTTCTACCTGGAAAATGTATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAGCAA
ATGCCACAAACCTGTGCAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGTTCAAAGGTTACCAGGG
AGACCTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACATGGAACCTGCCATGAACCCAAACAAATGCCA
ATGTCAAGAAGGTTGGCATGGAAGACACTGCAATAAAAGGTACGAAGCCAGCCTCATACATGCCCTGAGGCCAGC
AGGCGCCAGCTCAGGCAGCACAGCCTTCACTTAAAAGGCCGAGGAGCGGCGGGATCCACCTGAATCCAATTA
CATCTGG**TGA**ACTCCGACATCTGAAACGTTTTAAGTTACACCAAGTTCATAGCCTTTGTTAACCTTTTCATGTGTT
GAATGTTCAAATAATGTTCACTTACCTTAAGAATACTGGCCTGAATTTTATTAGCTTCATTATAAATCACTGAGC
TGATATTTACTCTTCTTTAAGTTTTCTAAGTACGTCTGTAGCATGATGGTATAGATTTTCTTGTTTCAGTGCT
TTGGGACAGATTTTATATTATGTCAATTGATCAGGTTAAATTTTTCAGTGTGTAGTTGGCAGATATTTTCAAAT
TACAATGCATTTATGGTGTCTGGGGGCAGGGGAACATCAGAAAGGTTAAATGGGCAAAATGCGTAAGTCACAA
GAATTTGGATGGTGCAGTTAATGTTGAAGTTACAGCATTTTCAGATTTTATTGTCAGATATTTAGATGTTTGTAC
ATTTTAAAAAATGCTCTTAATTTTAAACTCTCAATACAATATATTTTGACCTTACCATTATTCCAGAGATTCA
GTATTAACAAAAAATTAACACTGTGGTAGTGGCATTAAACAATATAATATATTCTAAACACAATGAAATAG
GGAATATAATGTATGAAGTTTTTGCAATGGCTTGAAGCAATATAATATATTGTAAACAAAACACAGCTCTTACCT
AATAAACATTTTATACTGTTTGTATGTATAAAATAAAGGTGCTGCTTTAGTTTTTGGAAAAA

308/550

FIGURE 308

MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESLYLWIDAHQARVLIGFEEDILIVSEGKM
APFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFSLRSLDKGIMADPTVNVPLLGT
VPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPPGGCRNG
GFCNERRICECPDGFHGPHEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCSTTCFN
GGTCFYPGKCICPPGLEGEQCEISKCPQPCRNGGKICIGSKCKCSKGYQGDLCSKPVCEPGCG
AHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAEERRDPPESNYIW

Important features:**Signal sequence:**

Amino acids 1-28

N-glycosylation sites:

Amino acids 88-92;245-249

Tyrosine kinase phosphorylation site:

Amino acids 370-378

N-myristoylation sites:

Amino acids 184-190;185-191;189-195;315-321

ATP/GTP-binding site motif A (P-loop):

Amino acids 285-293

EGF-like domain cysteine pattern signatures:

Amino acids 198-210;230-242;262-274;294-306;326-338

309/550

FIGURE 309

CCCACGCGTCCGGTCTCGCTCGCTCGCGCAGCGGCGGCAGCAGAGGTGCGGCACAGATGCGGG
TTAGACTGGCGGGGGAGGAGGCGGAGGAGGGAAGGAAGCTGCATGCATGAGACCCACAGACT
CTTGCAAGCTGGATGCCCTCTGTGGATGAAAGATGATCATGGAATGAACCCGAGCAATGGAG
ATGGATTTCTAGAGCAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCAGAGACTCTTGGCCG
TGATCCTGTGGTTTCAGCTGGCGCTGTGCTTCGGCCCTGCACAGCTCACGGGCGGGTTCGATG
ACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCGAGCGGAGGGG
TTTTCTTTGAAGGCTCTGTAGCCCGATTTCAGTCCAAGACGGATTCAAGCTGAAGGGCGCTA
CAAAGAGACTGTGTTGAAGCATTTTAATGGAACCCTAGGCTGGATCCCAAGTGATAATTCCA
TCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTCAATAACAAGACAT
ATAGACATGGAGAGAAGCTAATCATCACTTGTTCATGAAGGATTCAAGATCCGGTACCCCGACC
TACACAATATGGTTTCATTATGTCGCGATGATGGAACGTGGAATAATCTGCCCATCTGTCAAG
GCTGCCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGCTCCAGACCTCCT
TCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAACTTGATGGGTCTGCGT
ATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCACCCCGGTGCCTTGCTCTGGAAG
CCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTCGTCTGCCACCCGCGGCCTT
GTGAGCGCTACAACCACGGAAGTGTGGTGGAGTTTTACTGCGATCCTGGCTACAGCCTCACCA
GCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTCCTTCTTATCAAGTCTACTGCA
TCAAATCAGAGCAAACGTGGCCAGCACCCATGAGACCCTCCTGACCACGTGGAAGATTGTGG
CGTTCACGGCAACCAAGTGTGCTGCTGGTGTGCTGCTCGTCATCCTGGCCAGGATGTTCCAGA
CCAAGTTCAGGCCCCACTTTCCCCCAGGGGGCCTCCCCGGAGTTCAGCAGTGACCTGACT
TTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCCTATGACGAAGCTGTGAGTGGCGGCT
TGAGTGCTTAGGCCCCGGGTACATGGCCTCTGTGGGCCAGGGCTGCCCCTTACCCGTGGACG
ACCAGAGCCCCCAGCATACCCCGGCTCAGGGGACACGGACACAGGCCAGGGGAGTCAGAAA
CCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAAGTCTGTATTACCTCCCAGGTGCC
AAGAGAGCACCCACCCTGCTTCGGACAACCCTGACATAATTGCCAGCACGGCAGAGGAGGTGG
CATCCACCAGCCCAGGCATCCATCATGCCCAGTGGGTGTTGTTCTTAAGAACTGATTGATTA
AAAAATTTCCCAAAGTGTCTGAAGTGTCTCTTCAAATACATGTTGATCTGTGGAGTTGATTC
CTTTCCTTCTCTTGGTTTTAGACAAATGTAAACAAAGCTCTGATCCTTAAATTGCTATGCTG
ATAGAGTGGTGAGGGCTGGAAGCTTGATCAAGTCCTGTTTCTTCTTGACACAGACTGATTAAA
AATTAAAGNAAAAAA

310/550

FIGURE 310

MYHGMNPSNGDGFLEQQQQQQPQSPQRLLAVILWFQLALCFGPAQLTGGFDDLQVCADPGIP
ENGFRTPSGGVFFEGSVARFHCQDGFKLKGATKRLCLKHFNGTLGWIPSDNSICVQEDCRIPQ
IEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPICQGCLRPLASSNG
YVNISELQTSFPVGTVISYRCFPGFKLDGSAYLECLQNLIWSSSPPRCLALEAQVCPLPPMVS
HGDFVCHPRPCERYNHGTVVEFYCDPGYSLTSDYKYITCQYGEWFPSYQVYCIKSEQTWPSTH
ETLLTTWKIVAFTATSVLLVLLLVILARMEQTKFKAHFPPRGPPRSSSSDPDFVVVDGVPVML
PSYDEAVSGGLSALGPGYMASVGQGCPLPVDDQSPPAYPGSGD TDTGPGESETCDSVSGSSEL
LQSLYSPPRCQESTHPASDNPDIIASTAEVASTSPGIHHAHWVFLRN

Important features:**Signal sequence:**

amino acids 1-41

Transmembrane domain:

amino acids 325-344

N-glycosylation site.

amino acids 104-108, 134-138, 192-196

Casein kinase II phosphorylation site.amino acids 8-12, 146-150, 252-256, 270-274, 313-317, 362-366,
364-368, 380-384, 467-471, 468-472**N-myristoylation site.**amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424,
478-484**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 394-405

311/550

FIGURE 311

CAGCGCGTGGCCGGCGCCGCTGTGGGGACAGCATGAGCGGCGGTTGGATGGCGCAGGTTGGAG
CGTGGCGAACAGGGGCTCTGGGCCTGGCGCTGCTGCTGCTGCTCGGCCTCGGACTAGGCCTGG
AGGCCGCCGCGAGCCCGCTTTCCACCCCGACCTCTGCCCAGGCCGCGAGGCCCCAGCTCAGGCT
CGTGCCCAACCCACCAAGTTCCAGTGCCGCACCAAGTGGCTTATGCGTGCCCCCTCACCTGGCGCT
GCGACAGGGACTTGGACTGCAGCGATGGCAGCGATGAGGAGGAGTGCAGGATTGAGCCATGTA
CCCAGAAAGGGCAATGCCACCGCCCCCTGGCCTCCCCTGCCCCCTGCACCGGCGTCAGTGACT
GCTCTGGGGGAAGTGAACAAGAACTGCGCAACTGCAGCCGCCTGGCCTGCCTAGCAGGCGAGC
TCCGTTGCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCGACGGCCACCCAGACT
GTCCCGACTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGGAAGGGGATGCCA
CAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTACCTCTCTCAGGAATGCCACAACCATGG
GGCCCCCTGTGACCCTGGAGAGTGTCCCCCTCTGTGCGGAATGCCACATCCTCCTCTGCCGGAG
ACCAGTCTGGAAGCCCAACTGCCTATGGGGTTATTGCAGCTGCTGCGGTGCTCAGTGCAAGCC
TGGTCACCGCCACCCTCCTCCTTTTGTCTGGCTCCGAGCCCAGGAGCGCCTCCGCCCCACTGG
GGTTACTGGTGGCCATGAAGGAGTCCCTGCTGCTGTCAGAACAGAAGACCTCGCTGCCCTTGAG
GACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGGACAGGAGGAGAGCAGTGA
TGCGGATGGGTACCCGGGCACACCAGCCCTCAGAGACCTGAGTTCTTCTGGCCACGTGGAACC
TCGAACCCGAGCTCCTGCAGAAGTGGCCCTGGAGATTGAGGGTCCCTGGACACTCCCTATGGA
GATCCGGGGAGCTAGGATGGGGAACCTGCCACAGCCAGAACTGAGGGGCTGGCCCCAGGCAGC
TCCCAGGGGGTAGAACGGCCCTGTGCTTAAGACACTCCCTGCTGCCCCGTCTGAGGGTGGCGA
TTAAAGTTGCTTC

312/550

FIGURE 312

MSGGWMAQVGAWRTGALGLALLLLGLGLGLEAAASPLSTPTSAQAAGPSSGSCPPTKFQCRT
SGLCVPLTWRCRDLDCSDGSDEEEECRIEPC TQKGQC P P P PGLPCPCTGVSDCSGGTDKKLRN
CSRLACLAGELRCTLSDDCIPLTWRC DGHPCPDSSDELGCGTNEILPEGDATTMGPPVTLES
VTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAVLSASLVTATLLLLSW
LRAQERLRPLGLLVAMKESLLLSEQKTSLP

Important features:**Signal sequence:**

Amino acids 1-30

Transmembrane domain:

Amino acids 231-248

N-glycosylation sites:

Amino acids 126-130;195-199;213-217

Casein kinase II phosphorylation site.

amino acids 84-88, 140-144, 161-165, 218-222

N-myristoylation sites:Amino acids 3-9;10-16;26-32;30-36;112-118;166-172;212-218;
224-230;230-236;263-269**Prokaryotic membrane lipoprotein lipid attachment site:**

Amino acids 44-55

Leucine zipper pattern:

Amino acids 17-39

313/550

FIGURE 313

CGGACGCGTGGGCGTCCGGCGGTGCGCAGAGCCAGGAGGCGGAGGCGCGCGGGCCAGCCTGGGCCCCAGCCCACAC
CTTCACCAGGGCCCAGGAGCCACCATGTGGCGATGTCCACTGGGGCTACTGCTGTTGCTGCCGCTGGCTGGCCAC
TTGGCTCTGGGTGCCCAGCAGGGTCGTGGGCGCCGGGAGCTAGCACCGGGTCTGCACCTGCGGGGCATCCGGGAC
GCGGGAGGCCGGTACTGCCAGGAGCAGGACCTGTGCTGCCGCGGCCGTGCCGACGACTGTGCCCTGCCCTACCTG
GGCGCCATCTGTTACTGTGACCTCTTCTGCAACCGCACGGTCTCCGACTGCTGCCCTGACTTCTGGGACTTCTGC
CTCGGCGTGCCACCCCTTTTCCCCCGATCCAAGGATGTATGCATGGAGGTCGTATCTATCCAGTCTTGGAACG
TACTGGGACAACTGTAACCGTTGCACCTGCCAGGAGAACAGGCAGTGGCATGGTGGATCCAGACATGATCAAAGC
CATCAACCAGGGCAACTATGGCTGGCAGGCTGGGAACCACAGCGCCTTCTGGGCGATGACCCTGGATGAGGGCAT
TCGCTACCGCCTGGGCACCATCCGCCCCTCTCCTCGGTTCATGAACATGCATGAAATTTATACAGTGTGAACCC
AGGGGAGGTGCTTCCCACAGCCTTCGAGGCCTCTGAGAAGTGGCCCAACCTGATTCATGAGCCTCTTGACCAAGG
CAACTGTGCAGGCTCCTGGGCTTCTCCACAGCAGCTGTGGCATCCGATCGTGTCTCAATCCATTCTCTGGGACA
CATGAGCCTGTCTGTGCGCCCCAGAACCTGTGTCTGTGACACCCACCAGCAGAGGGCTGCCGCGGTGGGCG
TCTCGATGGTGCCTGGTGGTTCCTGCGTCGCGGAGGGGTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGA
ACGAGACGAGGCTGGCCCTGCGCCCCCTGTATGATGCACAGCCGAGCCATGGGTGCGGGCAAGCGCCAGGCCAC
TGCCCACTGCCCCAACAGCTATGTTAATAAATGACATCTACCAGGTCACTCCTGTCTACCGCCTCGGCTCCAA
CGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCATGGAGGTGCATGAGGACTTCTT
CCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGCCTTGGGAGGCCAGAGAGATACCGCCGGCATGGGAC
CCACTCAGTCAAGATCACAGGATGGGGAGAGGAGACGCTGCCAGATGGAAGGACGCTCAAATACTGGACTGCGGC
CAACTCCTGGGGCCCAGCCTGGGGCGAGAGGGGCCACTTCCGCATCGTGC GCGGCGTCAATGAGTGCGACATCGA
GAGCTTCGTGCTGGGCGTCTGGGGCCGCGTGGGCATGGAGGACATGGGTTCATCACTGAGGCTGCGGGCACCCACGC
GGGGTCCGGCCTGGGATCCAGGCTAAGGGCCGGCGGAAGAGGCCCAATGGGGCGGTGACCCAGCCTCGCCCGA
CAGAGCCCGGGGCGCAGGCGGGCGCCAGGGCGCTAATCCGCGCGGGTTCCGCTGACGCAGCGCCCCGCTGGG
AGCCGCGGGCAGGCGAGACTGGCGGAGCCCCAGACCTCCAGTGGGGACGGGGCAGGGCCTGGCTGGGAAGAG
CACAGCTGCAGATCCCAAGCCTCTGGCGCCCCACTCAAGACTACCAAAGCCAGGACACCTCAAGTCTCCAGCCC
CAATACCCACCCCAATCCCGTATCTTTTTTTTTTTTTTTAGACAGGGTCTTGCTCCGTTGCCAGGTTGGAG
TGCAGTGGCCCATCAGGGCTCACTGTAACCTCCGACTCCTGGGTTCAAGTGACCCTCCACCTCAGCCTCTCAAG
TAGCTGGGACTACAGGTGCACCACACACCTGGCTAATTTTTGTATTTTTGTAAAGAGGGGGGTCTCACTGTGT
TGCCAGGCTGGTTTGAACCTCCTGGGCTCAAGCGGTCCACCTGCCTCCGCCTCCCAAAGTGTGGGATTGCAGG
CATGAGCCACTGCACCCAGCCCTGTATTCTTATCTTCAGATATTATTTTTCTTTTCACTGTTTTAAATAAAA
CCAAAGTATTGATAAAAAAAA

314/550

FIGURE 314

MWRCPLGLLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQLCCRGRADDCA
LPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPFPPIQGCMHGGRIYPVLGTYWDNCNRCT
CQENRQWHGGSRHDQSHQPGQLWLAGWEPQRLLGHDGP

Important features:**N-glycosylation site.**

amino acids 78-82, 161-165

Casein kinase II phosphorylation site.

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300,
411-415

N-myristoylation site.

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230,
269-275, 378-384, 442-448

Amidation site.

amino acids 26-30, 318-322

Eukaryotic thiol (cysteine) proteases histidine active site.

amino acids 398-409

315/550

FIGURE 315

CGGACGCGTGGGCCCCCTGGTGGGCCCAGCAAG**ATG**GATCTACTGTGGATCCTGCCCTCCCTGT
GGCTTCTCCTGCTTGGGGGGCCTGCCTGCCTGAAGACCCAGGAACACCCCAGCTGCCCAGGAC
CCAGGGAAGTGGAGCCAGCAAAGTTGTCCTCCTGCCAGTTGTCCCGGAGCTCCAGGAAGTC
CTGGGGAGAAGGGAGCCCCAGGTCCTCAAGGGCCACCTGGACCACCAGGCAAGATGGGCCCCA
AGGGTGAGCCAGGCCCCAGAACTGCCGGGAGCTGTTGAGCCAGGGCGCCACCTTGAGCGGCT
GGTACCATCTGTGCCTACCTGAGGGCAGGGCCCTCCCAGTCTTTTGTGACATGGACACCGAGG
GGGGCGGCTGGCTGGTGTTCAGAGGCGCCAGGATGGTTCTGTGGATTTCTTCCGCTCTTGGT
CCTCCTACAGAGCAGGTTTTGGGAACCAAGAGTCTGAATTCTGGCTGGGAAATGAGAATTTGC
ACCAGCTTACTCTCCAGGGTAACTGGGAGCTGCGGGTAGAGCTGGAAGACTTTAATGGTAACC
GTACTTTCGCCCCACTATGCGACCTTCCGCCTCCTCGGTGAGGTAGACCACTACCAGCTGGCAC
TGGGCAAGTTCTCAGAGGGCACTGCAGGGGATTCCCTGAGCCTCCACAGTGGGAGGCCCTTTA
CCACCTATGACGCTGACCACGATTCAAGCAACAGCAACTGTGCAGTGATTGTCCACGGTGCCT
GGTGGTATGCATCCTGTTACCGATCAAATCTCAATGGTCGCTATGCAGTGTCTGAGGCTGCCG
CCCACAAATATGGCATTGACTGGGCCTCAGGCCGTGGTGTGGGCCACCCCTACCGCAGGGTTC
GGATGATGCTTCGAT**TAG**GGGCACTCTGGCAGCCAGTGCCCTTATCTCTCCTGTACAGCTTCCGG
ATCGTCAGCCACCTTGCCTTTGCCAACCACCTCTGCTTGCCTGTCCACATTTAAAAATAAAAT
CATTTTAGCCCTTTCA

316/550

FIGURE 316

MDLLWILPSLWLLLLGGPACLKTQEHPSCPGPRELEASKVVLLPSCPGAPGSPGEKGAPGPQG
PPGPPGKMGPKGEPGPRNCRELLSQGATLSGWYHLCLPEGRALPVFCMDMTEGGGWLVFQRRQ
DGSVDFFRSWSSYRAGFGNQESEFWLGNENLHQLTLQGNWELRVELEDFNGNRTFAHYATFRL
LGEVDHYQLALGKFSEGTAGDSLHSGRPFTTYDADHDSSNSNCAVIVHGAWWYASCYRSNL
NGRYAVSEAAAHKYGIDWASGRGVGHPYRRVRMMLR

Important features:**Signal peptide:**

Amino acids 1-16

N-glycosylation site:

Amino acids 178-182

Glycosaminoglycan attachment site:

Amino acids 272-276

Tyrosine kinase phosphorylation site:

Amino acids 188-197

N-myristoylation sites:

Amino acids 16-22;89-95;144-150;267-273

Fibrinogen beta and gamma chains C-terminal domain signature:

Amino acids 242-255

317/550

FIGURE 317

CCCAAGCCAGCCGAGCCGCCAGAGCCGCGGGCCGCGGGGGTGTGCGGGGCCCAACCCAGG**AT**
GCTCCCCCTGCGCCTCCTGCCTACCCGGGTCTCTACTGCTCTGGGCGCTGCTACTGTTGCTCTT
GGGATCAGCTTCTCCTCAGGATTCTGAAGAGCCCGACAGCTACACGGAATGCACAGATGGCTA
TGAGTGGGACCCAGACAGCCAGCACTGCCGGGATGTCAACGAGTGTCTGACCATCCCTGAGGC
CTGCAAGGGGGAAATGAAGTGCATCAACCACTACGGGGGCTACTTGTGCCTGCCCCGCTCCGC
TGCCGTCATCAACGACCTACATGGCGAGGGACCCCCGCCACCAGTGCCTCCCGCTCAACACCC
CAACCCCTGCCCCACCAGGCTATGAGCCCCGACGATCAGGACAGCTGTGTGGATGTGGACGAGTG
TGCCAGGGCCCTGCACGACTGTGCCCCAGCCAGGACTGCCATAACTTGCTGGCTCCTATCA
GTGCACCTGCCCTGATGGTTACCGCAAGATCGGGCCCCGAGTGTGTGGACATAGACGAGTGCCG
CTACCGCTACTGCCAGCACCGCTGCGTGAACCTGCCTGGCTCCTTCCGCTGCCAGTGCAGGCC
GGGCTTCCAGCTGGGGCCTAACAACCGCTCCTGTGTTGATGTGAACGAGTGTGACATGGGGGC
CCCATGCGAGCAGCGCTGCTTCAACTCCTATGGGACCTTCTGTGTGCTGCCACCAGGGCTA
TGAGCTGCATCGGGATGGCTTCTCCTGCAGTGATATTGATGAGTGTAGCTACTCCAGCTACCT
CTGTGAGTACCGCTGCGTCAACGAGCCAGGCCGTTTCTCCTGCCACTGCCACAGGGTTACCA
GCTGCTGGCCACACGCCCTCTGCCAAGACATTGATGAGTGTGAGTCTGGTGCGCACCAGTGCTC
CGAGGCCCAAACCTGTGTCAACTTCCATGGGGGCTACCGCTGCGTGGACACCAACCGCTGCGT
GGAGCCCTACATCCAGGTCTCTGAGAACCGCTGTCTCTGCCCCGGCCTCCAACCTCTATGTG
AGAGCAGCCTTCATCCATTGTGCACCGCTACATGACCATCACCTCGGAGCGGAGCGTGCCCG
TGACGTGTTCCAGATCCAGGCGACCTCCGTCTACCCCGGTGCCTACAATGCCTTTCAGATCCG
TGCTGGAAACTCGCAGGGGGACTTTTACATTAGGCAAATCAACAACGTCAGCGCCATGCTGGT
CCTCGCCCCGGCGGTGACGGGCCCCCGGGAGTACGTGCTGGACCTGGAGATGGTCACCATGAA
TTCCCTCATGAGCTACCGGGCCAGCTCTGTACTGAGGCTCACCGTCTTTGTAGGGGCCTACAC
CTTCT**TGA**GGAGCAGGAGGGAGCCACCCTCCCTGCAGCTACCCTAGCTGAGGAGCCTGTTGTGA
GGGGCAGAATGAGAAAGGCAATAAAGGGAGAAAGAAAGTCCTGGTGGCTGAGGTGGGCGGGTC
AACTGCAGGAAGCCTCAGGCTGGGGCAGGGTGGCACTTGGGGGGGCAGGCCAAGTTCACCTA
AATGGGGGTCTCTATATGTTTCAGGCCCAGGGGCCCCCATTGACAGGAGCTGGGAGCTCTGCAC
CACGAGCTTCAGTCACCCCGAGAGGAGAGGAGGTAACGAGGAGGGCGGACTCCAGGCCCCGGC
CCAGAGATTTGGACTTGGCTGGCTTGCAGGGGTCTAAGAACTCCACTCTGGACAGCGCCAG
GAGGCCCTGGGTTCATTCTTAACCTCTGCCTCAAACCTGTACATTTGGATAAGCCCTAGTAGTT
CCCTGGGCCTGTTTTTCTATAAAACGAGGCAACTGGAAAAAAAAAAAAA

318/550

FIGURE 318

MLPCASCLPGSLLLWALLLLLLLGSASPQDSEEPDSYTECTDGYEWDPSQHCRDVNECLTIPE
ACKGEMKCINHYGGYLCLPRSAAVINDLHGEGPPPPVPPAQHPNPCPPGYEPDDQDSCVDVDE
CAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRYCQHRCVNLPGSFRCQCE
PGFQLGPNNRSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYELHRDGFSCSDIDECSSYSSY
LCQYRCVNEPGRFSCHCPQGYQLLATRLCQDIDECESGAHQCSAQTCVNFHGGYRCVDTNRC
VEPYIQVSENRLCPASNPLCREQPSSIVHRYMTITSERVFPADVFIQATSVYPGAYNAFQI
RAGNSQGDIFYIRQINNVSAMLVLARPVTGPREYVLDLEMVTMNSLMSYRASSVLRLTVFVGAYTF

Important features:**Signal sequence:**

Amino acids 1-25

N-glycosylation sites:

Amino acids 198-202;394-398

N-myristoylation sites:Amino acids 76-82;145-151;182-188;222-228;290-296;305-311;
371-377;381-387**Aspartic acid and asparagine hydroxylation sites:**

amino acids 140-152;177-189;217-229;258-270

319/550

FIGURE 319

GCTGGGGAC**ATG**AGAGGCACACCGAAGACCCACCTCCTGGCCTTCTCCCTCCTCTGCCTCCTC
TCAAAGGTGCGTACCCAGCTGTGCCCAGACCATGTACCTGCCCCTGGCCACCTCCCCGATGC
CCGCTGGGAGTACCCCTGGTGCTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGGCGGCTG
GGGAGCCCTGCGACCAACTCCACGTCTGCGACGCCAGCCAGGGCCTGGTCTGCCAGCCCGGG
GCAGGACCCGGTGGCCGGGGGGCCCTGTGCCTCTTGCCAGAGGACGACAGCAGCTGTGAGGTG
AACGGCCGCCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCGCTGC
GAGGACGGCGGCTTACCTGCGTGCCGCTGTGCAGCGAGGATGTGCGGCTGCCAGCTGGGAC
TGCCCCACCCAGGAGGGTCGAGGTCTGGGCAAGTGCTGCCCTGAGTGGGTGTGCGGCCAA
GGAGGGGGACTGGGGACCCAGCCCCTTCCAGCCCAAGGACCCCAAGTTTCTGGCCTTGTCTCT
TCCCTGCCCCCTGGTGTCCCCTGCCAGAATGGAGCACGGCCTGGGGACCCTGCTCGACCACC
TGTGGGCTGGGCATGGCCACCCGGGTGTCCAACCAGAACCGCTTCTGCCGACTGGAGACCCAG
CGCCGCCTGTGCCTGTCCAGGCCCTGCCACCCCTCCAGGGGTCGCAGTCCACAAAACAGTGCC
TTC**TAG**AGCCGGGCTGGGAATGGGGACACGGTGTCCACCATCCCCAGCTGGTGGCCCTGTGCC
TGGGCCCTGGGCTGATGGAAGATGGTCCGTGCCAGGCCCTTGGCTGCAGGCAACACTTTAGC
TTGGGTCCACCATGCAGAACACCAATATTAACACGCTGCCTGGTCTGTCTGGATCCCGAGGTA
TGGCAGAGGTGCAAGACCTAGTCCCCTTTCTCTAACTCACTGCCTAGGAGGCTGGCCAAGGT
GTCCAGGGTCTCTAGCCCACTCCCTGCCTACACACACAGCCTATATCAAACATGCACACGGG
CGAGCTTTCTCTCCGACTTCCCCTGGGCAAGAGATGGGACAAGCAGTCCCTTAATATTGAGGC
TGCAGCAGGTGCTGGGCTGGACTGGCCATTTTTCTGGGGGTAGGATGAAGAGAAGGCACACAG
AGATTCTGGATCTCCTGCTGCCTTTTCTGGAGTTTGTAATAATTGTCCTGAATACAAGCCTAT
GCGTGA

320/550

FIGURE 320

MRGTPKTHLLAFSLLCLLSKVRTQLCPTPCTCPWPPPRCPLGVPLVLDGCGCCRVCAARRLGEP
CDQLHVCDASQGLVCQPGAGPGGRGALCLLAEDDSSCEVNGRLYREGETFQPHCSIRCRCEG
GFTCVPLCSEDVRLPSWDCPHPRRVEVLGKCCPEWVCQGGGLGTQPLPAQGPQFSGLVSSLP
PGVPCPEWSTAWGPCSTTCGLGMATRVSNQNRFCRLETQRRRLCLSRPCPPSRGRSPQNSAF

Important features:

Signal sequence:

Amino acids 1-23

N-myristoylation sites:

Amino acids 3-9;49-55;81-87;85-91;126-132;164-170;166-172;
167-173;183-189;209-215

Insulin-like growth factor binding proteins signature:

Amino acids 49-65

von Willebrand C1 domain:

Amino acids 107-124

Thrombospondin 1 Homology Block:

Amino acids 201-216

IGF binding protein site:

Amino acids 49-58

321/550

FIGURE 321

[illegible]

322/550

FIGURE 322

MMGLSLASAVLLASLLSLHLGTATRGSDISKTC CFQYSHKPLPWTWVRSYEFTSNSCSQRAVI
FTTKRGKKVCTHPRKKWVQKYISLLKTPKQL

Important features:

Signal peptide:

amino acids 1-23

N-myristoylation sites.

amino acids 3-9, 26-32

Amidation site.

amino acids 68-72

Small cytokines (intecrine/chemokine).

amino acids 23-88

323/550

FIGURE 323

ACCGAGCCGAGCGGACCGAAGGCGCGCCGAGATGAGGTGAGCAAGAGGATGCTGGCGGGGGCGTGAGGAGCA
TGCCAGCCCCCTCCTGGCCTGCTGGCAGCCCATCCTCCTGCTGGTGTGGGCTCAGTGCTGTGAGGCTCGGCCA
CGGGTGCCCGCCCGCTGCGAGTGCTCCGCCAGGACCGCGCTGTGCTGTGCCACCGCAAGTGCTTTGTGGCAG
TCCCCGAGGGCATCCCCACCGAGACGCGCCTGCTGGACCTAGGCAAGAACCGCATCAAAACGCTCAACCAGGACG
AGTTCGCCAGCTTCCCGACCTGGAGGAGCTGGAGCTCAACGAGAACATCGTGAGCGCCGTGGAGCCCCGGCGCCT
TCAACAACCTCTTCAACCTCCGGACGCTGGGTCTCCGACGCAACCGCCTGAAGCTCATCCGCTAGGCGTCTTCA
CTGGCCTCAGCAACCTGACCAAGCAGGACATCAGCGAGAACAAGATCGTTATCCTACTGGACTACATGTTTCAGG
ACCTGTACAACCTCAAGTCACTGGAGGTTGGCGACAATGACCTCGTCTACATCTCTACCGCGCCTTCAGCGGCC
TCAACAGCCTGGAGCAGCTGACGCTGGAGAAATGCAACCTGACCTCCATCCCCACCGAGGCGCTGTCCACCTGC
ACGGCCTCATCGTCTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGCTGTACC
GACTCAAGGTCTTGAGATCTCCCACTGGCCCTACTTGGACACCATGACACCAACTGCCTCTACGGCTCAACC
TGAGCTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCCCTACCTGGCGTCCGCCACCTAGTCTATCTCC
GCTTCCTCAACCTCTCCTACAACCCCATCAGCACCATTGAGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCAGG
AGATCCAGCTGGTGGCGGGCAGCTGGCCGTGGTGGAGCCCTATGCCTTCCGCGGCCTCAACTACCTGCGCGTGC
TCAATGTCTCTGGCAACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGGAGACACTCA
TCCTGGACTCCAACCCGCTGGCCTGCGACTGTCCGCTCCTGTGGGTGTTCCGGCGCCGCTGGCGGCTCAACTTCA
ACGGGCAGCAGCCACGTGCGCCACGCCCAGTTTGTCCAGGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTAC
TGCCCAACTACTTCACTGCGCGCGCGCCGCATCCGGGACCGCAAGGCCAGCAGGTGTTTGTGGACGAGGGCC
ACACGGTGCAGTTTGTGTGCCGGCCGATGGCGACCCGCGCCGCCCATCCTCTGGCTCTACCCCGAAAGCACC
TGGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCCCTGATGGCAGCTGGAGGTGCGCTACGCCCAGGTAC
AGGACAACGGCACGTACCTGTGCATCGCGGCAACGCGGGCGGCAACGACTCCATGCCCCGCCACCTGCATGTGC
GCAGCTACTCGCCGACTGGCCCCATCAGCCCAACAAGACCTTCGCTTTCATCTCCAACCAGCCGGGCGAGGGAG
AGGCCAACAGCACCCGCGCCACTGTGCCTTTCCTTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGCT
TCATCTCTTTCCTGGGCGTCTCCTCTTCTGCCTGGTGTGCTGTTTCTCTGGAGCCGGGGCAAGGGCAACACAA
AGCACAACATCGAGATCGAGTATGTGCCCCGAAAGTCGGACGCAGGCATCAGCTCCGCGACGCGCCCCGCAAGT
TCAACATGAAGATGATATGAGGCCGGGGCGGGGGCAGGGACCCCCGGGCGGCCGGCAGGGGAAGGGGCCGTGGT
CGCCACCTGCTCACTCTCCAGTCCTTCCACCTCCTCCCTACCCTTCTACACACGTTCTCTTCTCCCTCCCGCC
TCCGTCCCCTGCTGCCCCCGCCAGCCCTCACCACCTGCCCTCCTTCTACCAGGACCTCAGAAGCCCAGACCTGG
GGACCCACCTACACAGGGGCATTGACAGACTGGAGTTGAAAGCCGACGAACCGACACGCGGCAGAGTCAATAAT
TCAATAAAAAAGTTACGAACCTTCTCTGTAACCTGGGTTTCAATAATTATGGATTTTATGAAAACCTGAAATAA
TAAAAAGAGAAAAAACTAAAAAAAAAAAAAAAAAAAAA

324/550

FIGURE 324

MQVSKRMLAGGVRSMPSPILLACWQPILLVLGSLVSGSATGCPPRCECSAQDRAVLCHRKCFVAVPEGIPTETRL
LDLGKNRIKTLNQDEFASFPHLEELNENIVSAVEPGAFFNNLFLNRLTLGLRSNRLKLIPLGVFTGLSNLTKQDI
SENKIVILLDYMFDLYNLKSLEVGDNLDVYISHRAFSGLSLEQLTLEKCNLTSIPTEALSHLHGLIVLRLRHL
NINAIRDYSFKRLYLKVLKLEISHWPYLDTMTPNCLYGLNLTSLSITHCNLTAVPYLAVRHLVYLRFNLNSYNPIS
TIEGSMHELLRLQEIQLVGGQLAVVEPYAFRGLNYLRVLNVSGNQLTTLEESVFHSGVGNLETILDSNPLACDC
RLLWVFRRRLNLFNRQQPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQVFVDEGHTVQFVCRADG
DPPPAIWLSPRKHLVSAKSNRGLTVFPDGTLEVRYAQVQDNGTYLCIAANAGGND SMPAHLHVRSYSPDWPHQP
NKTFAFISNQPGEGEANSTRATVPFPFDIKTLIIATTMGFISFLGVVLFCLVLLFLWSRGKGN TKH NIEIEYVPR
KSDAGISSADAPRKFNMKMI .

Important features:**Signal sequence:**

amino acids 1-41

Transmembrane domain:

amino acids 556-578

N-glycosylation site.amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345, 492-496,
505-509, 526-530, 542-546**Casein kinase II phosphorylation site.**

amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353, 607-611

Tyrosine kinase phosphorylation site.

amino acids 590-598

N-myristoylation site.amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143, 262-268, 320-326,
344-350, 359-365, 493-499, 503-509, 605-611**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 32-43

325/550

FIGURE 325

CCCACGCGTCCGCCCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGAG
GCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGAGG
AGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAAGGAGAGAGGGAGGAGGAGGAG
GAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAACGGAG
AGGAGGTGTGGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGAGTAGG
AAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGGAAAGAC
ACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTGGCTGCTT
TGGCATTGTTGGGAACTGGGACTCCCTGTGGGGAGGAGAGGAAAGCTGGAAGTCTGGAGGGAC
AGGGTCCCAGAAGGAGGGGACAGAGGAGCTGAGAGAGGGGGGCAGGGCGTTGGGCAGGGGTCC
CTCGGAGGCCTCCTGGGGATGGGGGCTGCAGCTCGTCTGAGCGCCCCCTCGAGCGCTGGTACTC
TGGGCTGCACTGGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCCGAGGACTGGTGGAGC
TACAAGGATAATCTCCAGGGAACTTCGTGCCAGGGCCTCCTTTCTGGGGCCTGGTGAATGCA
GCGTGGAGTCTGTGTGCTGTGGGGAAGCGGCAGAGCCCCGTGGATGTGGAGCTGAAGAGGGTT
CTTTATGACCCCTTTCTGCCCCATTAAGGCTCAGCACTGGAGGAGAGAAGCTCCGGGGAACC
TTGTACAACACCGGCCGACATGTCTCCTTCCTGCCTGCACCCCGACCTGTGGTCAATGTGTCT
GGAGGTCCCCTCCTTTACAGCCACCGACTCAGTGAAGTGCAGGCTGCTGTTTGGAGCTCGCGAC
GGAGCCGGCTCGGAACATCAGATCAACCACCGGGCTTCTCTGCTGAGGTGCAGCTCATTCAC
TTCAACCAGGAAGTCTACGGGAATTTACAGCGCTGCCTCCCGCGGCCCCAATGGCCTGGCCATT
CTCAGCCTCTTTGTCAACGTTGCCAGTACCTCTAACCATTCTCAGTCGCCTCCTTAACCGC
GACACCATCACTCGCATCTCCTACAAGAATGATGCCTACTTTCTTCAAGACCTGAGCCTGGAG
CTCCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGGGCTCTCTCAGCACCCCGCCCTGC
TCCGAGACTGTCACCTGGATCCTCATTGACCGGGCCCTCAATATCACCTCCCTTCAGATGCAC
TCCCTGAGACTCCTGAGCCAGAACTCCTCCATCTCAGATCTTCCAGAGCCTCAGCGGTAACAGC
CGGCCCCTGCAGCCCTTGGCCACAGGGCACTGAGGGGCAACAGGGACCCCGGCACCCCGAG
AGGCGCTGCCGAGGCCCCAACTACCGCCTGCATGTGGATGGTGTCCCCCATGGTCGCTGAGAC
TCCCCTTCGAGGATTGCACCCGCCCGTCCTAAGCCTCCCCACAAGGCGAGGGGAGTTACCCCT
AAAACAAAGCTATTAAAGGGACAGAATACTTA

326/550

FIGURE 326

MGAAARLSAPRALVLWAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLCA
VGKRQSPVDVELKRVLYDPFLPPLRLSTGGELRGTYLNTGRHVSFLPAPRPVVNVSGGPLY
SHRLSELRLLEFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSLEVN
VASTSNPFLSRLLRDTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSTPPCSETVTW
ILIDRALNITSLQMHSRLRLSQNPFSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPERRCRGP
NYRLHVDGVPHGR

Important features:**Signal peptide:**

Amino acids 1-23

Transmembrane domain:

Amino acids 177-199

N-glycosylation sites:

Amino acids 118-122;170-174;260-264

Eukaryotic-type carbonic anhydrases proteins:

Amino acids 222-271;128-165;45-93

327/550

FIGURE 327

GGACTAATCTGTGGGAGCAGTTTATTCCAGTATCACCCAGGGTGCAGCCACACCAGGACTGTGTTGAAGGGTGT
TTTTTCTTTTAAATGTAATACCTCCTCATCTTTTCTTCTTACACAGTGTCTGAGAACATTTACATTATAGATAA
GTAGTACATGGTGGATAACTTCTACTTTTAGGAGGACTACTCTCTTCTGACAGTCTTAGACTGGTCTTCTACACT
AAGACACCATGAAGGAGTATGTGCTCCTATTATTCTGGCTTTGTGCTCTGCCAAACCCTTCTTTAGCCCTTCAC
ACATCGCACTGAAGAATATGATGCTGAAGGATATGGAAGACACAGATGATGATGATGATGATGATGATGATGATG
ATGATGATGAGGACAACTCTCTTTTCCAACAAGAGAGCCAAGAAGCCATTTTTTTCCATTTGATCTGTTTCCAA
TGTGTCCATTTGGATGTCAGTGCTATTACAGAGTTGTACATTGCTCAGATTTAGGTTTGACCTCAGTCCCAACCA
ACATTCCATTTGATACTCGAATGCTTGATCTTCAAAACAATAAAATTAAGGAAATCAAAGAAAATGATTTTAAAG
GACTCACTTCACTTTATGGTCTGATCCTGAACAACAACAGCTAACGAAGATTCACCCAAAGCCTTTCTAACCA
CAAAGAAGTTGCGAAGGCTGTATCTGTCCACAATCACTAAGTGAAATACCACTTAATCTTCCCAAATCATTAG
CAGAACTCAGAATTCATGAAAATAAAGTTAAGAAAATACAAAAGGACACATTCAAAGGAATGAATGCTTTACAG
TTTTGGAAATGAGTGCAAACCCTCTTGATAATAATGGGATAGAGCCAGGGGCATTTGAAGGGGTGACGGTGTTC
ATATCAGAATTGCAGAAGCAAACTGACCTCAGTTCCTAAAGGCTTACCACCAACTTTATTGGAGCTTCACCTAG
ATTATAATAAAATTTCAACAGTGGAACCTTGAGGATTTTAAACGATACAAAGAACTACAAAGGCTGGGCCTAGGAA
ACAACAAAATCACAGATATCGAAAATGGGAGTCTTGCTAACATACCACGTGTGAGAGAAAATACATTTGGAAAACA
ATAAACTAAAAAAATCCCTTCAGGATTACCAGAGTTGAAATACCTCCAGATAATCTTCCTTCATTCTAATTCAA
TTGCAAGAGTGGGAGTAAATGACTTCTGTCCAACAGTGCCAAAGATGAAGAAATCTTTATACAGTGAATAAGTT
TATTCAACAACCCGGTGAAATACTGGGAAATGCAACCTGCAACATTTGTTGTGTTTTGAGCAGAATGAGTGTTT
AGCTTGGGAACCTTTGGAATGTAATAAATAGTAATTGGTAATGTCCATTTAATATAAGATTCAAAAATCCCTACAT
TTGGAATACTTGAACCTCTATTAATAATGGTAGTATTATATATACAAGCAAATATCTATTCTCAAGTGGTAAGTCC
ACTGACTTATTTTATGACAAGAAATTTCAACGGAATTTTGCCAACTATTGATACATAAGGGGTTGAGAGAAAACA
AGCATCTATTGCAGTTTCCTTTTGGGTACAAATGATCTTACATAAATCTCATGCTTGACCATTCCCTTCTTCAT
AACAAAAAGTAAGATATTCGGTATTTAACACTTTGTTATCAAGCACATTTTAAAAAGAACTGTACTGTAAATGG
AATGCTTGACTTAGCAAAATTTGTGCTCTTTCATTTGCTGTTAGAAAAACAGAATTAACAAAGACAGTAATGTGA
AGAGTGCATTACACTATTCTTATTCTTTAGTAACTTGGGTAGTACTGTAATATTTTAAATCATCTTAAAGTATGA
TTTGATATAATCTTATTGAAATTACCTTATCATGTCTTAGAGCCCGTCTTTATGTTTAAAACTAATTTCTTAAAA
TAAAGCCTTCAGTAAATGTTCAATACCAACTTGATAAATGCTACTCATAAGAGCTGGTTTGGGGCTATAGCATAT
GCTTTTTTTTTTTTAAATTATTACCTGATTTAAAAATCTCTGTAAAAACGTGTAGTGTTCATAAAATCTGTAAC
CGCATTTTAAATGATCCGCTATTATAAGCTTTTAAATAGCATGAAAATGTTAGGCTATATAACATTGCCACTTCAA
CTCTAAGGAATATTTTTGAGATATCCCTTTGGAAGACCTTGCTTGGAAGAGCCTGGACACTAACAAATCTACACC
AAATTGTCTCTTCAAATACGTATGGACTGGATAACTCTGAGAAACACATCTAGTATAACTGAATAAGCAGAGCAT
CAAATTAACAGACAGAAACCGAAAGCTCTATATAAATGCTCAGAGTTCTTTATGTATTTCTTATTGGCATTCAA
CATATGTAAATCAGAAAACAGGGAAATTTTCATTAAAAATATTGGTTTGAAAT

328/550

FIGURE 328

MKEYVLLLFLALCSAKPFFSPSHIALKNMMLKDMEDTDDDDDDDDDDDDDEDNSLFPTREPRS
HFFPFDLFPMCPFGCQCYSRVVHCSDLGLTSVPTNIPFDTRMLDLQNNKIKEIKENDFKGLTS
LYGLILNNNKLTKIHPKAFLTTKKLRRLYLSHNQLSEIPLNLPKSLAELRIHENKVKKIQKDT
FKGMNALHVLEMSANPLDNNGIEPGAFEGVTVFHIRIAEAKLTSVPKGLPPTLLELHLDYNKI
STVELEDFKRYKELQRLGLGNNKITDIENGLANIPRVREIHLENNKLKKIPSGLPPELKYLQI
IFLHSNSIARVGVNDFCPTVPKMKKSLYSAISLFNNPVKYWEMQPATFRCVLSRMSVQLGNFGM

Important features:**Signal sequence.**

amino acids 1-15

N-glycosylation site.

amino acids 281-285

N-myristoylation sites.

amino acids 129-135, 210-216, 214-220, 237-243, 270-276, 282-288

Leucine zipper pattern.

amino acids 154-176

329/550

FIGURE 329

GGGGTCTCCCTCAGGGCCGGGAGGCACAGCGGTCCCTGCTTGCTGAAGGGCTGGATGTACGCA
TCCGCAGGTTCCCGCGGACTTGGGGGCGCCCGCTGAGCCCCGGCGCCCGCAGAAGACTTGTGT
TTGCCTCCTGCAGCCTCAACCCGGAGGGCAGCGAGGGCCTACCACCATGATCACTGGTGTGTT
CAGCATGCGCTTGTGGACCCAGTGGGCGTCCTGACCTCGCTGGCGTACTGCCTGCACCAGCG
GCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGTCGACCGCAGCCTGCT
GAAGTTGAAAATGGTGCAGGTCGTGTTTCGACACGGGGCTCGGAGTCCTCTCAAGCCGCTCCC
GCTGGAGGAGCAGGTAGAGTGGAACCCCCAGCTATTAGAGGTCCCACCCCAAACCTCAGTTTGA
TTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCTCAATACCA
TGAGACCACCCTGAAGGGGGGCATGTTTGCTGGGCAGCTGACCAAGGTGGGCATGCAGCAAAT
GTTTGCCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCTTTCTTTACCAAC
CTTCAACCCACAGGAGGTCTTTATTGTTCCACTAACATTTTTTCGGAATCTGGAGTCCACCCG
TTGTTTGCTGGCTGGGCTTTTCCAGTGTGAGAAAGAAGGACCCATCATCATCCACACTGATGA
AGCAGATTCAGAAGTCTTGATCCCAACTACCAAAGCTGCTGGAGCCTGAGGCAGAGAACCAG
AGGCCGGAGGCAGACTGCCTCTTTACAGCCAGGAATCTCAGAGGATTTGAAAAAGGTGAAGGA
CAGGATGGGCATTGACAGTAGTGATAAAGTGGAATCTTTCATCCTCCTGGACAACGTGGCTGC
CGAGCAGGCACACAACCTCCCAAGCTGCCCCATGCTGAAGAGATTTGCACGGATGATCGAACA
GAGAGCTGTGGACACATCCTTGATACATACTGCCCAAGGAAGACAGGGAAAGTCTTCAGATGGC
AGTAGGCCCATTCCTCCACATCCTAGAGAGCAACCTGCTGAAAGCCATGGACTCTGCCACTGC
CCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATGTGACCTTCATACCGCTCTT
AATGACCCTGGGGATTTTTGACCACAAATGGCCACCGTTTGCTGTTGACCTGACCATGGAAC
TTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTATTACCACGGGAAGGAGCAGGT
GCCGAGAGGTGGCCCTGATGGGCTCTGCCCGCTGGACATGTTCTTGAATGCCATGTCAGTTTA
TACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAAACCTCAGGTGATGGAAGTTGGAAA
TGAAGAGTAACTGATTTATAAAAGCAGGATGTGTTGATTTTAAATAAAGTGCCTTTATACAATG

330/550

FIGURE 330

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLMVQVVFRHGARSPLKPLPLEEQVE
WNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQLTKVGMQQMFA LGERLRKNYVEDIPFL
SPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKEGP IIIHTDEADSEVLYPNYQSCWSLRQRTGRROTASLQ
PGISED LKKVKDRMGIDSSDKVDFILLDNVAAEQAHNLPSCPMLKR FARMIEQRAVDTSLYILPKEDRESLQMA
VGPFLHILES NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLMTLGIFDHKWPPFAVDLTMELYQHLESKEWFVQ
LYYHGKEQVPRGCPDGLCLDMFLNAMS VYTLSPEKYHALCSQTQVMEVGNEE

Important features:**Signal sequence:**

amino acids 1-23

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 218-222

Casein kinase II phosphorylation site.

amino acids 87-91, 104-108, 320-324

Tyrosine kinase phosphorylation site.

amino acids 280-288

N-myristoylation site.

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

Amidation site.

amino acids 216-220

Leucine zipper pattern.

amino acids 10-32

Histidine acid phosphatases phosphohistidine signature.

amino acids 50-65

331/550

FIGURE 331

CGAGGGCTTTTCCGGCTCCGGAATGGCACATGTGGGAATCCCAGTCTTGTTGGCTACAACATTTTCCCTTTCCT
AACAAATTCTAACAGCTGTTCTAACAGCTAGTGATCAGGGGTTCTTCTTGCTGGAGAAGAAAGGGCTGAGGGCAG
AGCAGGGCACTCTCACTCAGGGTGACCAGCTCCTTGCCCTCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGAT
GCAGCGGAGTGAGGTGATGGAAGTCTAAAATAGGAAGGAATTTTGTGTGCAATATCAGACTCTGGGAGCAGTTGA
CCTGGAGAGCCTGGGGGAGGGCCTGCCTAACAAAGCTTTCAAAAAACAGGAGCGACTTCCACTGGGCTGGGATAAG
ACGTGCCGGTAGGATAGGGAAGACTGGGTTTAGTCCTAATATCAAATTGACTGGCTGGGTGAACCTCAACAGCCT
TTTAACCTCTCTGGGAGATGAAACGATGGCTTAAGGGGCCAGAAATAGAGATGCTTTGTAAAATAAAATTTTAA
AAAAAGCAAGTATTTTATAGCATAAAGGCTAGAGACCAAAATAGATAACAGGATTCCCTGAACATTCTAAGAGG
GAGAAAGTATGTTAAAAATAGAAAAACCAAAATGCAGAAGGAGGAGACTCACAGAGCTAAACCAGGATGGGGACC
CTGGGTGTCAGGCCAGCCTCTTTGCTCCTCCCGGAAATTATTTTGGTCTGACCACTCTGCCTTGTTGTTTGCAGAA
TCATGTGAGGGCCAACCGGGGAAGGTGGAGCAGATGAGCACACACAGGAGCCGTCTCCTCACCGCCGCCCTCTC
AGCATGGAACAGAGGCAGCCCTGGCCCCGGGCCCTGGAGGTGGACAGCCGCTCTGTGGTCTGCTCTCAGTGGTC
TGGGTGCTGCTGGCCCCCAGCAGCCGGCATGCCTCAGTTCAGCACCTTCCACTCTGAGAATCGTGACTGGACC
TTCAACCACTTGACCGTCCACCAAGGGACGGGGCCGTCTATGTGGGGCCATCAACCGGTCTATAAGCTGACA
GGCAACCTGACCATCCAGGTGGCTCATAAGACAGGGCCAGAAGAGGACAACAAGTCTCGTTACCCGCCCTCATC
GTGCAGCCCTGCAGCGAAGTGCTCACCCCTACCAACAATGTCAACAAGCTGCTCATCATGACTACTCTGAGAAC
CGCTGCTGGCTGTGGGAGCCTCTACCAGGGGGTCTGCAAGCTGCTGCGGCTGGATGACCTCTTCATCCTGGTG
GAGCCATCCACAAGAAGGAGCACTACCTGTCCAGTGTCAACAAGACGGGCACCATGTACGGGGTGATTGTGCGC
TCTGAGGGTGAGGATGGCAAGCTCTTCATCGGCACGGCTGTGGATGGGAAGCAGGATTACTTCCCGACCTGTCC
AGCCGGAAGCTGCCCCGAGACCTGAGTCCTCAGCCATGCTCGACTATGAGCTACACAGCGATTTTGTCTCCTCT
CTCATCAAGATCCCTTCAGACACCTGGCCCTGGTCTCCCACTTTGACATCTTCTACATCTACGGCTTTGCTAGT
GGGGGCTTTGTCTACTTTCTCACTGTCCAGCCGAGACCCCTGAGGGTGTGGCCATCAACTCCGCTGGAGACCTC
TTCTACACCTCACGCATCGTGGGCTCTGCAAGGATGACCCCAAGTTCCACTCATACGTGTCCCTGCCCTTCGGC
TGCACCCGGGCGGGGTGGAATACCGCCTCCTGCAGGCTGCTTACCTGGCCAAGCCTGGGGACTCACTGGCCCAG
GCCTTCAATATCACCAGCCAGGACGATGTACTCTTGCCATCTTCTCAAAGGGCAGAAGCAGTATCACCACCCG
CCCGATGACTCTGCCCTGTGTGCCTTCCCTATCCGGGCCATCAACTTGCAGATCAAGGAGCGCCTGCAGTCCTGC
TACCAGGGCGAGGGCAACCTGGAGCTCAACTGGCTGCTGGGGAAGGACGTCCAGTGCACGAAGGCGCCTGTCCCC
ATCGATGATAACTTCTGTGGACTGGACATCAACCAGCCCCTGGGAGGCTCAACTCCAGTGGAGGGCCTGACCCTG
TACACCACCAGCAGGGACCGCATGACCTCTGTGGCCTCCTACGTTTACAACGGCTACAGCGTGGTTTTTGTGGGG
ACTAAGAGTGGAAGCTGAAAAAGGTAAGAGTCTATGAGTTCAGATGCTCCAATGCCATTACCTCCTCAGCAAA
GAGTCCCTCTTGAAGGTAGCTATTGGTGGAGATTTAACTATAGGCAACTTTATTTTCTTGGGGAACAAAGGTGA
AATGGGGAGGTAAGAAGGGTTAATTTTGTGACTTAGCTTCTAGCTACTTCCCTCCAGCCATCAGTCATTGGGTAT
GTAAGGAATGCAAGCGTATTTCAATATTTCCCAAACCTTAAGAAAAACTTTAAGAAGGTACATCTGCAAAAGCAAA

332/550

FIGURE 332

MGTLGQASLFAPPGNFYFWS DHSA LCF AESCEGQPGKVEQMSTHRSRLLTAAPLSMEQRQPWPR
ALEVDSRSVVLLSVVWVLLAPPAAGMPQFSTFHSEN RDWTFNHLTVHQGTGAVYVGAINRVYK
LTGNLTIQVAHKTGPEEDNKSRY PPLIVQPCSEVLTLTNNVNKLLIIDYSENRL LACGS LYQG
VCKLLRLDDLFI LVEPSHKKEHYLSSVNKTGTMYGVIVRSEGEDGKLFIGTAVD GKD YFPTL
SSRKLPRDP ESSAMLDYELHSD FVSSLIKIPSDTLALVSHFDIFYIYG FASGGFVYFLT VQPE
TPEGVAINSAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKPGDSL AQ
AFNITSQDDVLF AIFSKGQKQYHHPDDSALCAFP IRAINLQIKERLQSCYQGE GNLELNWLL
GKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASYVYNGYSVVFVGT
KSGKLKKVRVYEFRC SNAIHLLSKESLLEGSYWWRFNYRQLYFLGEQR

Important features:**Signal sequence:**

amino acids 1-32

Transmembrane domain:

amino acids 71-87

N-glycosylation site.

amino acids 130-134, 145-149, 217-221, 381-385

Casein kinase II phosphorylation site.amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387,
384-388, 471-475, 481-485, 530-534**N-myristoylation site.**

amino acids 220-226, 319-325, 353-359, 460-466, 503-509

333/550

FIGURE 333

GCTGAGTCTGCTGCTCCTGCTGCTGCTGCTCCAGCCTGTAACCTGTGCCTACACCACGCCAGG
CCCCCCCAGAGCCCTCACCACGCTGGGCGCCCCCAGAGCCCACACCATGCCGGGCACCTACGC
TCCCTCGACCACACTCAGTAGTCCCAGCACCCAGGGCCTGCAAGAGCAGGCACGGGCCCTGAT
GCGGGACTTCCCGCTCGTGGACGGCCACAACGACCTGCCCTGGTCCTAAGGCAGGTTTACCA
GAAAGGGCTACAGGATGTTAACCTGCGCAATTTAGCTACGGCCAGACCAGCCTGGACAGGCT
TAGAGATGGCCTCGTGGGCGCCAGTTCTGGTCAGCCTATGTGCCATGCCAGACCCAGGACCG
GGATGCCCTGCGCCTCACCCTGGAGCAGATTGACCTCATACGCCGCATGTGTGCCTCCTATTC
TGAGCTGGAGCTTGTGACCTCGGCTAAAGCTCTGAACGACACTCAGAAATTGGCCTGCCTCAT
CGGTGTAGAGGGTGGCCACTCGCTGGACAATAGCCTCTCCATCTTACGTACCTTCTACATGCT
GGGAGTGGCGTACCTGACGCTCACCCACACCTGCAACACACCCTGGGCAGAGAGCTCCGCTAA
GGGCGTCCACTCCTTCTACAACAACATCAGCGGGCTGACTGACTTTGGTGAGAAGGTGGTGGC
AGAAATGAACCGCCTGGGCATGATGGTAGACTTATCCCATGTCTCAGATGCTGTGGCACGGCG
GGCCCTGGAAGTGTACAGGCACCTGTGATCTTCTCCCACTCGGCTGCCCGGGGTGTGTGCAA
CAGTGCTCGGAATGTTCTGATGACATCCTGCAGCTTCTGAAGAAGAACGGTGGCGTCGTGAT
GGTGTCTTTGTCCATGGGAGTAATACAGTGCAACCCATCAGCCAATGTGTCCACTGTGGCAGA
TCACTTCGACCACATCAAGGCTGTCATTGGATCCAAGTTCATCGGGATTGGTGGAGATTATGA
TGGGGCCGGCAAATTCCTCAGGGGCTGGAAGACGTGTCCACATACCCGGTCCTGATAGAGGA
GTTGCTGAGTCGTGGCTGGAGTGAGGAAGAGCTTCAGGGTGTCTTCGTGGAAACCTGCTGCG
GGTCTTCAGACAAGTGGAAAAGGTACAGGAAGAAAACAAATGGCAAAGCCCCTTGGAGGACAA
GTTCCCGGATGAGCAGCTGAGCAGTTCCTGCCACTCCGACCTCTCACGTCTGCGTCAGAGACA
GAGTCTGACTTCAGGCCAGGAACTCACTGAGATTCCCATACTGGACAGCCAAGTTACCAGC
CAAGTGGTCAGTCTCAGAGTCCTCCCCCACATGGCCCCAGTCCTTGCAGTTGTGGCCACCTT
CCCAGTCCTTATTCTGTGGCTCTGATGACCCAGTTAGTCCTGCCAGATGTCACTGTAGCAAGC
CACAGACACCCACAAAGTTCCCCTGTTGTGCAGGCACAAATATTTCTGAAATAAATGTTTT
GGACATAG

334/550

FIGURE 334

MPGTYAPSTTLSSPSTQGLQEQARALMRDFPLVDGHNDLPLVLRQVYQKGLQDVNLRNFSYGQ
TSLDRLRDGLVGAQFWSAYVPCQTQDRDALRLTLEQIDLIIRMCASYSELELVTSAKALNDTQ
KLACLIGVEGGHSLDNSLSILRTFYMLGVRYLTLTHTCNTPWAESSAKGVHSFYNNISGLTDF
GEKVVAEMNRLGMMVDLSHVSDAVARRALEVSQAPVIFSHSAARGVCNSARNVPDDILQLLKK
NGGVVMVSLSMGVIQCNPSANVSTVADHFDHIKAVIGSKFIGIGGDYDGAGKFPQGLEDVSTY
PVLIEELLSRGWSEELQGVLRGNLLRVFRQVEKVQEENKWQSPLEDKFPDEQLSSSCHSDLS
RLRQRQSLTSGQELTEIPIHWTAKLPAKWSVSESSPHMAPVLAVVATFPVLILWL

Important features:**N-glycosylation sites.**

amino acids 58-62, 123-127, 182-186, 273-277

N-myristoylation sites.

amino acids 72-78, 133-139, 234-240, 264-270, 334-340, 389-395

Renal dipeptidase active site.

amino acids 134-157

335/550

FIGURE 335

CCCAGAAAGTTCAAGGGCCCCCGGCTCCTGCGCTCCTGCCGCCGGGACCCTCGACCTCCTCAG
AGCAGCCGGCTGCCGCCCCGGGAAGATGGCGGAGGAGGAGCCGCCACCGCCTCCTCCTGCTGCT
GCTGCGCTACCTGGTGGTGCCTTGGGCTATCATAAGGCCTATGGGTTTTCTGCCCCAAAAGA
CCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTTAGCCTGCAAACCCCAAAGAA
GACTGTTTTCTCCAGATTAGAGTGGAAGAACTGGGTTCGGAGTGTCTCCTTTGTCTACTATCA
ACAGACTCTTCAAGGTGATTTTAAAAATCGAGCTGAGATGATAGATTTCAATATCCGGATCAA
AAATGTGACAAGAAGTGATGCGGGGAAATATCGTTGTGAAGTTAGTGCCCCATCTGAGCAAGG
CCAAAACCTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTTCCATCATG
TGAAGTACCCTCTTCTGCTCTGAGTGGAAGTGTGGTAGAGCTACGATGTCAAGACAAAGAAGG
GAATCCAGCTCCTGAATACACATGGTTTAAGGATGGCATCCGTTTGCTAGAAAATCCCAGACT
TGGCTCCCAAAGCACCAACAGCTCATAACAATGAATACAAAACTGGAAGTCTGCAATTTAA
TACTGTTTTCCAACTGGACACTGGAGAATATTCTGTGAAGCCCGCAATTCTGTTGGATATCG
CAGGTGTCCTGGGAAACGAATGCAAGTAGATGATCTCAACATAAGTGGCATCATAGCAGCCGT
AGTAGTTGTGGCCTTAGTGATTTCCGTTTGTGGCCTTGGTGTATGCTATGCTCAGAGGAAAGG
CTACTTTTCAAAGAAACCTCCTTCCAGAAGAGTAATTCTTCATCTAAAGCCACGACAATGAG
TGAAAATGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGAAGGCCGCGGGCGGGGATC
ACGAGGTCAGGAGTTCTAGACCAGTCTGGCCAATATGGTGAAACCCCATCTCTACTAAAATAC
AAAAATTAGCTGGGCATGGTGGCATGTGCCTGCAGTTCCAGCTGCTTGGGAGACAGGAGAATC
ACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATCACGCCACTGCAGTCCAGCCTGGG
TAACAGAGCAAGATTCCATCTCAAAAAATAAAATAAATAAATAAATAAATAAATAAATAAATAA
TGTTAGAAATCTTACAATAAATATAGCTTGATATTC

336/550

FIGURE 336

MARRSRHRLLLLLLRYLVVALGYHKAYGFSAPKDQQVVTAVEYQEAILACKTPKKTVSSRLEW
KKLGRSVSVFVYYQQTLOGDFKNRAEMIDFNIRIKNVTRSDAGKYRCEVSAPSEQGQNLEEDTV
TLEVLVAPAVPSCEVPSSALSGTVVELRCQDKEGNPAPEYTWFKDGIRLLENPRLGSQSTNSS
YTMNTKTGTLQFNTVSKLDTGEYSCEARNSVGYYRRCPGKRMQVDDLNISGIIAAVVVVALVIS
VCGLGVCYAQRKGYFSKETSFQKSNSSSKATTMSENVQWLTPVIPALWKAAAGGSRGQEF

Important features:**Signal peptide:**

amino acids 1-20

Transmembrane domain:

amino acids 130-144, 238-258

N-glycosylation site.

amino acids 98-102, 187-191, 236-240, 277-281

Casein kinase II phosphorylation site.

amino acids 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

N-myristoylation site.

amino acids 182-188, 239-245, 255-261, 257-263, 305-311

Amidation site.

amino acids 226-230

337/550

FIGURE 337

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTGCGCGAGCCTCGG
CACCTGCAGGTCCGTGCGTCCCGCGGCTGGCGCCCCTGACTCCGTCCCGGCCAGGGAGGGCCA
TGATTTCCCTCCCGGGGCCCCCTGGTGACCAACTTGCTGCGGTTTTTGTTCTTGGGGCTGAGTG
CCCTCGCGCCCCCTCGCGGGCCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAGGCGG
TGGAGGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGACGGGGAGGTGTCTTCATCCC
AGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATCAGGTGT
TGTCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCCATGCCCT
CCCGGAACCTGTCCCTGCGGCTGGAGGTCTCCAGGAGAAAGACTCTGGCCCCTACAGCTGCT
CCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCTTAGAACTCA
ATGTACTGGTTCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCATGTGGGGGCAA
ACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCAGTGGGATCGGC
AGCTTCCATCCTTCCAGACTTTCTTTGCCACCAGCATTAGATGTCATCCGTGGGTCTTTAAGCC
TCACCAACCTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCACAAATGAGGTGGGCA
CTGCCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGCAGTGGTTGCTGGAG
CTGTTGTGGGTACCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCCTCTTGTACCACCGCC
GGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCCATTGCTCCCCGGACCC
TGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCTTTCCTCTGTACCTCCG
CACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCATTGACCCCCACGCCAGTC
TCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCACGACAGATGGGGCCCACCCTCAACCAA
TATCCCCCATCCCTGGTGGGGTTTCTTCCTCTGGCTTGAGCCGCATGGGTGCTGTGCCTGTGA
TGGTGCCTGCCCAGAGTCAAGCTGGCTCTCTGGTAT**TG**ATGACCCCACTCATTGGCTAAAG
GATTTGGGGTCTCTCCTTCCCTATAAGGGTCACCTCTAGCACAGAGGCCTGAGTCATGGGAAAG
AGTCACACTCCTGACCCTTAGTACTCTGCCCCACCTCTCTTTACTGTGGGAAAACCATCTCA
GTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGAAGTGGATCTGGAATTGGGAGGAGC
CTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTGCTGAAATTAGCTACTACCAAGAGT
GAGGGGCAGAGACTTCCAGTCACTGAGTCTCCAGGCCCCCTTGATCTGTACCCACCCCTAT
CTAACACCACCCCTTGGCTCCCACTCCAGCTCCCTGTATTGATATAACCTGTCAGGCTGGCTTG
GTTAGGTTTTTACTGGGGCAGAGGATAGGGAATCTCTTATTAAACTAACATGAAATATGTGTT
GTTTTCATTTGCAAATTTAAATAAAGATACATAATGTTTGTATGAAAAA

338/550

FIGURE 338

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSSS
QPWEVPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKGDSGPYSC
SVNVQDKQGKSRGHSIKTLELNLVPPAPPSCRLQGVPHVGANVTLSQCSPRSKPAVQYQWDR
QLPSFQTFFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNTLEVSTGPGAADVAG
AVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLWPWKSSDTISKNGTLSSVTS
ARALRPPHGP RP GAL TPTPSLSSQALPSPRLPTTDGAHPQPISPIPGGVSSSGLSRMGAVPV
MVPAQSQAGSLV

Important features:**Signal peptide:**

amino acids 1-29

Transmembrane domain:

amino acids 245-267

N-glycosylation site.

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

N-myristoylation site.amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262,
262-268, 308-314, 363-369, 364-370**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 164-175

339/550

FIGURE 339

GCGAGAACCTTTGACGCGCACAAACTACGGGGACGATTTCTGATTGATTTTGGCGCTTTTCGATCCACCCTCCT
CCCTTCTCATGGGACTTTGGGGACAAAGCGTCCCGACCGCTCGAGCGCTCGAGCAGGGCGCTATCCAGGAGCCA
GGACAGCGTCGGGAACCGAGACCATGGCTCCTGGACCCCAAGATCCTTAAGTTCGTCTTCATCGTCGGGTTT
TGCTGCCGGTCCGGGTTGACTCTGCCACCATCCCCCGGCAGGACGAAGTTCCCCAGCAGACAGTGGCCCCCAGC
AACAGAGGCGCAGCCTCAAGGAGGAGGAGTGTCCAGCAGGATCTCATAGATCAGAATATACTGGAGCCTGTAACC
CGTGCACAGAGGGTGTGGATTACACCATTTGCTTCCAACAATTTGCCTTCTTGCCTGCTATGTACAGTTTGTAAT
CAGGTCAAACAAATAAAAGTTCTGTACCACGACAGAGACACCGTGTGTAGTGTGAAAAAGGAGCTTCCAGG
ATAAAAACTCCCCTGAGATGTGCCGGACGTGTAGAACAGGGTGTCCCAGAGGGATGGTCAAGGTCAGTAATTGTA
CGCCCCGGAGTGACATCAAGTGCAAAATGAATCAGCTGCCAGTTCCACTGGGAAAACCCAGCAGCGGAGGAGA
CAGTGACCACCATCCTGGGATGCTTGCCTCTCCCTATCACTACCTTATCATCATAGTGGTTTTAGTCATCATTT
TAGCTGTGGTTGTGGTTGGCTTTTCATGTGCGAAGAAATTCATTTCTTACCTCAAAGGCATCTGCTCAGGTGGTG
GAGGAGTCCCGAACGTGTGCACAGAGTCCTTTCCGGCGCGGTTTCATGTCTTACAGGTTCTTGGGGCGGAGG
ACAATGCCCGCAACGAGACCCCTGAGTAACAGATACTTGACGCCACCCAGGTCCTCTGAGCAGGAAATCCAAGTC
AGGAGCTGGCAGAGCTAACAGGTGTGACTGTAGAGTCGCCAGAGGAGCCACAGCGTCTGCTGGAACAGGCAGAA
CTGAAGGGTGTGAGGAGGAGGCTGCTGGTTCCAGTGAATGACGCTGACTCCGCTGACATCAGCACCTTGCTGG
ATGCCCTCGGCAACACTGGAAGAAGGACATGCAAAGGAAACAATTCAGGACCAACTGGTGGGCTCCGAAAAGCTCT
TTTATGAAGAAGATGAGGCAGGCTCTGCTACGTCCTGCCTGTGAAGAATCTCTTCAGGAAACCAGAGCTTCCCT
CATTTACCTTTTCTCCTACAAAGGGAAGCAGCCTGGAAGAACAGTCCAGTACTTGACCCATGCCCAACAACT
CTACTATCCAATATGGGCGAGCTTACCAATGGTCTAGAACTTTGTTAACGCACTTGGAGTAATTTTATGAAT
ACTGCGTGTGATAAGCAAACGGGAGAAATTTATATCAGATTCTTGGCTGCATAGTTATACGATTGTGTATTAAGG
GTCGTTTTAGGCCACATGCGGTGGCTCATGCCTGTAATCCAGCACTTTGATAGGCTGAGGCAGGTGGATTGCTT
GAGCTCGGGAGTTTGAGACAGCCTCATCAACACAGTGAAGTCCATCTCAATTTAAAAAGAAAAAAGTGGTTT
TAGGATGTCATTTCTTGCAGTTCTTCATCATGAGACAAGTCTTTTTTCTGCTTCTTATATTGCAAGCTCCATCT
CTACTGGTGTGTGCATTTAATGACATCTAACTACAGATGCCGACAGCCACAATGCTTTGCCTTATAGTTTTTTA
ACTTTAGAACGGGATTATCTTGTATTACCTGTATTTTCAGTTTCCGATATTTTGACTTAATGATGAGATTATC
AAGACGTAGCCCTATGCTAAGTCATGAGCATATGGACTTACGAGGGTTCGACTTAGAGTTTGAGCTTTAAGATA
GGATTATTGGGGCTTACCCCCACCTTAATTAGAGAAACATTTATATTGCTTACTACTGTAGGCTGTACATCTCTT
TTCCGATTTTTGTATAATGATGTAAACATGGAATACTTTAGGAAATGCACTTATTAGGCTGTTTACATGGGTTG
CCTGGATACAAATCAGCAGTCAAAAATGACTAAAAATATAACTAGTGACGGAGGGAGAAATCCTCCCTCTGTGGG
AGGCACTTACTGCATTCCAGTTCTCCCTCCTGCGCCCTGAGACTGGACCAGGGTTTGATGGCTGGCAGCTTCTCA
AGGGGCAGCTTGTCTTACTTGTTAATTTTAGAGGTATATAGCCATATTTATTTATAAATAAATATTTATTTATTT
ATTTATAAGTAGATGTTTACATATGCCAGGATTTTGAAGAGCCTGGTATCTTTGGGAAGCCATGTGTCTGGTTT
GTCGTGCTGGGACAGTCATGGGACTGCATCTTCCGACTTGTCCACAGCAGATGAGGACAGTGAGAATTAAGTTAG
ATCCGAGACTGCGAAGAGCTTCTCTTTCAAGCGCCATTACAGTTGAACGTTAGTGAATCTTGAGCCTCATTTGGG
CTCAGGGCAGAGCAGGTGTTTATCTGCCCCGCATCTGCCATGGCATCAAGAGGGAAGAGTGGACGGTGTCTGGG
AATGGTGTGAAATGGTTGCCGACTCAGGCATGGATGGGCCCCCTCTCGCTTCTGGTGGTCTGTGAAGTGAAGTCCCT
GGGATGCCTTTTAGGGCAGAGATTCCTGAGCTGCGTTTTAGGGTACAGATTCCCTGTTTGAGGAGCTTGGCCCCCT
CTGTAAGCATCTGACTCATCTCAGAGATATCAATTCCTTAAACACTGTGACAACGGGATCTAAATAGGCTGACACA
TTTGTCTTGTGTACGTTCCATTATTTTATTTAAAAACCTCAGTAATCGTTTTAGCTTCTTTCCAGCAAACTCT
TCTCCACAGTAGCCAGTCGTGGTAGGATAAATTACGGATATAGTCATTCTAGGGGTTTCAGTCTTTTCCATCTC
AAGGCATTGTGTGTTTTGTTCCGGGACTGGTTTGGCTGGGACAAAGTTAGAAGTGCCTGAAGTTCGCACATTGAG
ATTGTTGTGTCATGGAGTTTTAGGAGGGGATGGCTTTCCGGTCTTCGCACTTCCATCCTCTCCCACTTCCATC
TGGCGTCCACACCTTGTCCCTGCACCTTCTGGATGACACAGGGTGTGCTGCTGCTCCTAGTCTTTGCTTTGCTG
GGCCTTCTGTGCAGGAGACTTGGTCTCAAAGCTCAGAGAGAGCCAGTCCGGTCCAGCTCCTTTGTCCCTTCCCTC
AGAGGCCTTCCCTGAAGATGCATCTAGACTACCAGCCTTATCAGTGTTTAAGCTTATTCCTTTAACATAAGCTTC
CTGACAACATGAAATGTTGGGGTTTTTGGCGTTGGTTGATTGTTTAGGTTTTGCTTTATACCCGGGCCAAAT
AGCACATAACACCTGTTTATATGAAATACTCATATGTTTATGACCAAAATAAATATGAAACCTCATRTTAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

340/550

FIGURE 340

MGLWGQSVPTASSARAGRYPGARTASGTRPWLLDPKILKFVVFIVAVLLPVRVDSATIPRQDEVPPQQTVA PQQQR
RSLKEEECPAGSHRSEYTGACNPCTEGVDYTIASNNLPSCLLCTVCKSGQTNKSSCTTTRDTVCQCEKGSFQDKN
SPEMCRTRCTGCPRGMVKVSNCTPRSDIKCKNES AASSTGKTPAAEETVTTILGMLASPYHYLIIIVVLVIILAV
VVVGFSRKKFISYLGKICSGGGGGPERVHRVLFRRRSCPSRVPGAEDNARNETLSNRYLQPTQVSEQEIQGQEL
AELTGVTVESPEEPQRLLEQAEAGCQRRRLVPVNDADSADISTLLDASATLEEGHAKETIQDQLVGSEKLFYE
EDEAGSATSC

Important features:**Transmembrane domains:**

amino acids 35-52, 208-230

N-glycosylation sites.

amino acids 127-131, 182-186, 277-281

Glycosaminoglycan attachment site.

amino acids 245-249

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 260-264

N-myristoylation sites.

amino acids 21-27, 86-92, 102-108, 161-167, 242-248, 270-276, 297-303, 380-386

ATP/GTP-binding site motif A (P-loop).

amino acids 185-193

TNFR/NGFR cysteine-rich region.

amino acids 99-139

341/550

FIGURE 341

GCCTCTGAATTGTTGGGCAGTCTGGCAGTGGAGCTCTCCCCGGTCTGACAGCCACTCCAGAGG
CCATGCTTTCGTTTCTTGCCAGATTTGGCTTTCAGCTTCCTGTTAATTCTGGCTTTGGGCCAGG
CAGTCCAATTTCAAGAATATGTCTTTCTCCAATTTCTGGGCTTAGATAAGGCGCCTTCACCCC
AGAAGTTCCAACCTGTGCCTTATATCTTGAAGAAAATTTCCAGGATCGCGAGGCAGCAGCGA
CCACTGGGGTCTCCCGAGACTTATGCTACGTAAAGGAGCTGGGCGTCCGCGGAATGTACTTC
GCTTTCCTCCAGACCAAGGTTTCTTTCTTTACCCAAAGAAAATTTCCCAAGCTTCCTCCTGCC
TGCAGAAGCTCCTCTACTTTAACCTGTCTGCCATCAAAGAAAGGGAACAGTTGACATTGGCCC
AGCTGGGCCTGGACTTGGGGCCCAATTCTTACTATAACCTGGGACCAGAGCTGGAAGTGGCTC
TGTTCTGTTTCAGGAGCCTCATGTGTGGGGCCAGACCACCCCTAAGCCAGGTAAAATGTTTG
TGTTGCGGTCAGTCCCATGGCCACAAGGTGCTGTTCACTTCAACCTGCTGGATGTAGCTAAGG
ATTGGAATGACAACCCCCGGAAAAATTTCTGGGTTATTCTCTGGAGATACTGGTCAAAGAAGATA
GAGACTCAGGGGTGAATTTTCAGCCTGAAGACACCTGTGCCAGACTAAGATGCTCCCTTCATG
CTTCCCTGCTGGTGGTGACTCTCAACCCTGATCAGTGCCACCCTTCTCGGAAAAGGAGAGCAG
CCATCCCTGTCCCAAGCTTCTTGTAAAGAACCTCTGCCACCGTCACCAGCTATTCATTAAGT
TCCGGGACCTGGGTTGGCACAAGTGGATCATTTGCCCCAAGGGGTTTCATGGCAAATTACTGCC
ATGGAGAGTGTCCCTTCTCACTGACCATCTCTCTCAACAGCTCCAATTATGCTTTCATGCAAG
CCCTGATGCATGCCGTTGACCCAGAGATCCCCCAGGCTGTGTGTATCCCCACCAAGCTGTCTC
CCATTTCCATGCTCTACCAGGACAATAATGACAATGTCATTCTACGACATTATGAAGACATGG
TAGTCGATGAATGTGGGTGTGGGTAGGATGTCAGAAATGGGAATAGAAGGAGTGTTCTTAGGG
TAAATCTTTTAATAAACTACCTATCTGGTTTATGACCACTTAGATCGAAATGTC

342/550

FIGURE 342

MLRFLPDLAFSFLLILALGQAVQFQEYVFLQFLGLDKAPSPQKFQVPYILKKIFQDREAAAT
TGVSRDLCYVKELGVRGNVLRFLPDQGFFLYPKKISQASSCLQKLLYFNLSAIKEREQLTLAQ
LGLDLGPNSYYNLGPELELALFLVQEPHVWGQTPKPGKMFVLRVWPWPQGAVHFNLLDVAKD
WNDNPRKNFGLFLEILVKEDRDSGVNFQPEDTCARLRCSLHASLLVVTLPDQCHPSRKRRAA
IPVPKLSCKNLCHRHQLFINFRDLGWHKWIIAPKGFMANYPCHGECPFSLTISLNSSNYAFMQA
LMHAVDPEIPQAVCIPTKLSPISMLYQDNNDNVILRHYEDMVVDECGCG

Important features:**Signal peptide:**

amino acids 1-21

N-glycosylation sites.

amino acids 112-116, 306-310

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 96-100

N-myristoylation site.

amino acids 77-83

TGF-beta family proteins.

amino acids 264-299, 327-341, 345-364

343/550

FIGURE 343

CCCACGCGTCCGGCCTTCTCTCTGGACTTTGCATTTCATTCTTTTCATTGACAACTGACTTTTTTTATTTCT
TTTTTCCATCTCTGGGCCAGCTTGGGATCCTAGGCCGCCCTGGGAAGACATTTGTGTTTTACACACATAAGGAT
CTGTGTTTTGGGGTTTCTTCTCCTCCCCTGACATTGGCATTGCTTAGTGTTGTGTGGGGAGGGAGACCAGTGG
GCTCAGTGCTTGTCTGCACTTATCTGCCTAGGTACATCGAAGTCTTTTGACCTCCATACAGTGATTATGCCTGTC
ATCGCTGGTGGTATCCTGGCGGCCTTGTCTGCTGATAGTTGTCTGCTCTGTCTTTACTTCAAAATACACAAC
GCGCTAAAAGCTGCAAAGGAACCTGAAGCTGTGGCTGTAAAAAATCACAACCCAGACAAGGTGTGTGGGCCAAG
AACAGCCAGGCCAAAACCATTGCCACGGAGTCTTGTCTGCCCTGCAGTGCTGTGAAGGATATAGAATGTGTGCC
AGTTTTGATTCCCTGCCACCTTGTCTGTTGCGACATAAATGAGGGCCTCTGAGTTAGGAAAGGCTCCCTTCTCAA
GCAGAGCCCTGAAGACTTCAATGATGTCAATGAGGCCACCTGTTTGTGATGTGCAGGCACAGAAAGGCACAG
CTCCCCATCAGTTTTCATGGAAAATAACTCAGTGCCTGCTGGGAACCAGCTGCTGGAGATCCCTACAGAGAGCTTC
CACTGGGGGCAACCCTTCCAGGAAGGAGTTGGGGAGAGAGAACCCCTCACTGTGGGGAATGCTGATAAACAGTCA
CACAGTGTCTTATTCTCACAAAATCTACCCCTTGCCTGGCTGGAAGTGCAGTTTCCCTGGAGGTGTCCAGAAA
GCTGATGTAACACAGAGCCCTATAAAAAGCTGTGGTCTTTAAGGCTGCCAGCGCCTTGCCAAATGGAGCTTGTA
AGAAGGCTCATGCCATTGACCCCTCTTAATTCTCTCTGTTTGGCGGAGCTGACAATGGCGGAGGCTGAAGGCAAT
GCAAGCTGCACAGTCAGTCTAGGGGGTGCCAATATGGCAGAGACCCACAAAGCCATGATCCTGCAACTCAATCCC
AGTGAGAAGTGCACCTGGACAATAGAAAGACCAGAAAACAAAAGCATCAGAAATATCTTTTCTATGTCCAGCTT
GATCCAGATGGAAGCTGTGAAAGTGAAACATTAAAGTCTTTGACGGAACCTCCAGCAATGGGCCTCTGTAGGG
CAAGTCTGCAGTAAAAACGACTATGTTCTGTATTGTAATCATCATCCAGTACATTGACGTTTCAAATAGTTACT
GACTCAGCAAGAATTCAAAGAACTGTCTTTGTCTTCTACTACTTCTTCTCTCTCTAACATCTCTATTCCAACTGT
GGCGGTTACCTGGATACCTTGGAAAGGATCCTTACCAGCCCCAATTACCCAAAGCCGCATCCTGAGCTGGCTTAT
TGTGTGTGGCACATACAAGTGGAGAAAGATTACAAGATAAACTAACTTCAAAGAGATTTTCTAGAAATAGAC
AAACAGTGCAAATTTGATTTTCTTGCCATCTATGATGGCCCTCCACCAACTCTGGCCTGATTGGACAAGTCTGT
GGCCGTGTGACTCCACCTTCGAATCGTCATCAAACCTCTCTGACTGTCTGTGTTGTCTACAGATTATGCCAATTCT
TACCGGGGATTTTCTGCTTCTTACACCTCAATTTATGCAGAAAACATCAACACTACATCTTTAACTTGCTCTTCT
GACAGGATGAGAGTTATTATAAGCAAATCCTACCTAGAGGCTTTTAACTCTAATGGGAATAACTTGCAACTAAAA
GACCCAACTTGACAGACCAAATTTATCAAATGTTGTGGAATTTTCTGTCCCTCTTAATGGATGTGGTACAATCAGA
AAGGTAGAAGATCAGTCAATTACTTACACCAATATAATCACCTTTTCTGCATCCTCAACTTCTGAAGTGATCACC
CGTCAGAAACAACCTCCAGATTATTGTGAAGTGTGAAATGGGACATAATTCTACAGTGGAGATAATATACATAACA
GAAGATGATGTAATACAAAGTCAAATGCACTGGGCAAATATAACACCAGCATGGCTCTTTTGAATCCAATTCA
TTTGAAAAGACTTACTTGAATCACCATATTATGTGGATTTGAACCAAACCTCTTTTGTCAAGTTAGTCTGCAC
ACCTCAGATCCAAATTTGGTGGTGTCTTGTGATACCTGTAGAGCCTCTCCACCTCTGACTTTGCATCTCCAACC
TACGACCTAATCAAGAGTGGATGTAGTTCGAGATGAACTTGTAAGGTGTATCCCTTATTGGACACTATGGGAGA
TTCCAGTTTAAATGCCTTTAAATCTTGAGAAGTATGAGCTCTGTGTATCTGCAGTGTAAAGTTTTGATATGTGAT
AGCAGTGACCACAGTCTCGCTGCAATCAAGGTTGTGTCTCCAGAAGCAAACGAGACATTTCTTCATATAAATGG
AAAACAGATTCCATCATAGGACCCATTCTGTCTGAAAAGGGATCGAAGTGAAGTGGCAATTTCAGGATTCAGCAT
GAAACACATGCGGAAGAACTCCAAACCAGCCTTTCAACAGTGTGCATCTGTTTCTTTCATGGTTCTAGCTCTG
AATGTGGTGAATGTAGCGACAATCACAGTGAGGCATTTTGTAAATCAACGGGCAGACTACAAATACCAGAAGCTG
CAGAATATTAATAACAGGTCCAACCCTAAGTGAGACATGTTTCTCCAGGATGCCAAAGGAAATGCTACCTCGT
GGCTACACATATTATGAATAAATGAGGAAGGGCCTGAAAGTGACACACAGGCCTGCATGTAAAAAA

344/550

FIGURE 344

MELVRRRLMPLTLLILSCLAELTMAEAEGNASCTVSLGGANMAETHKAMILQLNPSENCTWTIE
RPENKSIRIIFSYVQLDPDGSCSENIKVFDTSSNGPLLQGVCCKNDYVPVFESSSSTLTFTQ
IVTDSARIQRTVFVFYFFSPNISIPNCGGYLDLTLEGSFTSPNYPKPHPELAYCVWHIQVEKD
YKIKLNFKEIFLEIDKQCKFDFLAIYDGPSTNSGLIGQVCGRVPTPTFESSNSLTVVLSTDYA
NSYRGFSASYTSIYAENINTTSLTCSSDRMRVVIISKSYLEAFNSNGNNLQLKDPTCRPKLSNV
VEFSVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQKQLQIIVKCEMGNSTVEIIYI
TEDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQTLFVQVSLHTSDPNLVVFLDT
CRASPTSDFASPTYDLIKSGCSRDETCKVYPLFGHYGRFQFNAFKFLRSMSSVYLQCKVLICD
SSDHQSRCNQGCVSRSKRDISSYKWKTD SIIGPIRLKRDRSASGNSGFQHETHAEETPNQPFN
SVHLFSFMVLALNVTVATITVRHFVNQRADYKYQKLQNY

Important features:**Signal sequence:**

amino acids 1-24

Transmembrane domain:

amino acids 571-586

N-glycosylation site.amino acids 29-33, 57-61, 67-71, 148-152, 271-275, 370-374,
394-398, 419-423**Casein kinase II phosphorylation site.**amino acids 22-26, 108-112, 289-293, 348-352, 371-375, 379-383,
408-412, 463-467, 520-524, 556-560**Tyrosine kinase phosphorylation site.**

amino acids 172-180, 407-415, 407-416, 519-528

N-myristoylation site.

amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

345/550

FIGURE 345

TGGGGGCCCCCAGGCTCGCGCGTGGAGCGAAGCAGCATGGGCAGTCGGTGCGCGCTGGCCCTGGCGGTGCTCTC
GGCCTTGCTGTGTAGGTCTGGAGCTCTGGGGTGTTCGAAGTGAAGCTGCAGGAGTTCGTCAACAAGAGGGGGCT
GCTGGGGAACCGCAATTGCTGCGCGGGGGCGCGGGGCCACCGCGTGCGCCTGCCGGACCTTCTCCGCGTGTG
CCTCAAGCACTACCAGGCCAGCGTGTCCCCGAGCCGCCCTGCACCTACGGCAGCGCCGTACCCCCGTGCTGGG
CGTCGACTCCTTCAGTCTGCCCCGACGGCGGGGGCGCGACTCCGCGTTCAGCAACCCCATCCGCTTCCCTTCGG
CTTCACCTGGCCGGGCACCTTCTCTCTGATTATGAAGCTCTCCACACAGATTCTCTGATGACCTCGCAACAGA
AAACCCAGAAAGACTCATCAGCCGCTGGCCACCCAGAGGCACCTGACGGTGGGCGAGGAGTGGTCCCAGGACCT
GCACAGCAGCGGCCGACCGGACCTCAAGTACTCCTACCGCTTCGTGTGTGACGAACACTACTACGGAGAGGGCTG
CTCCGTTTTCTGCCGTCCCCGGGACGATGCCCTTCGGGCCACTTCACCTGTGGGGAGCGTGGGGAGAAAGTGTGCAA
CCCTGGCTGGAAAGGGCCCTACTGCACAGAGCCGATCTGCCCTGCCCTGGATGTGATGAGCAGCATGGATTTTGTGA
CAAACAGGGGAATGCAAGTGCAGAGTGGGCTGGCAGGGCCGGTACTGTGACGAGTGTATCCGCTATCCAGGCTG
TCTCCATGGCACCTGCCAGCAGCCCTGGCAGTGAAGTGCAGGAAGGGCTGGGGGGGCTTTTCTGCAACCAGGA
CCTGAAGTACTGCACACACCATAAGCCCTGCAAGAATGGAGCCACCTGCACCAACACGGGGCCAGGGGAGCTACAG
TTGCTCTTGCCGGCCTGGGTACACAGGTGCCACCTGCGAGCTGGGGATTGACGAGTGTGACCCAGCCCTTGTAA
GAACGGAGGGAGCTGCACGGATCTCGAGAACAGCTACTCCTGTACCTGCCCCACCGGCTTCTACGGCAAAATCTG
TGAATTGAGTGCCATGACCTGTGCGGACGGCCCTTGCTTTAACGGGGGTGGTGTCTCAGACAGCCCGATGGAGG
GTACAGCTGCCGCTGCCCGTGGGCTACTCCGGCTTCAACTGTGAGAAGAAATTTGACTACTGCAGCTCTTCACC
CTGTTCTAATGGTGCCAAGTGTGTGGACCTCGGTGATGCCCTACCTGTGCCGCTGCCAGGCCGGCTTCTCGGGGAG
GCACTGTGACGACAACGTGGAGCACTGCGCCTCCTCCCCGTGCGCCAACGGGGGCACCTGCGCGGATGGCGTGAA
CGACTTCTCTGCACCTGCCCGCTGGCTACACGGGCAGGAAGTGCAGTCCCCCGTCAAGCAGGTGCGAGCACGC
ACCCTGCCACAATGGGGCCACCTGCCACGAGAGGGGGCCACCGCTATGTGTGCGAGTGTGCCGAGGCTACGGGGG
TCCCAACTGCCAGTTCCTGCTCCCCGAGCTGCCCCGGGGCCAGCGGTGGTGGACCTCACTGAGAAGCTAGAGGG
CCAGGGCGGGCCATTCCCTGGGTGGCCGTGTGCGCCGGGGTCACTTGTCTCATGCTGCTGCTGGGCTGTGC
CGCTGTGGTGGTCTGCGTCCGGCTGAGGCTGCAGAAGCACCGGCCCCAGCCGACCCCTGCCGGGGGGAGACGGA
GACCATGAACAACCTGCCCAACTGCCAGCGTGAGAAGGACATCTCAGTCAGCATCATCGGGGCCACGCAGATCAA
GAACACCAACAAGAAGGCGGACTTCCACGGGGACCACAGCGCCGACAAGATGGCTTCAAGGCCGCTACCCAGC
GGTGGACTATAACCTCGTGCAGGACCTCAAGGGTGACGACACCGCCGTCAAGGACCGGCACAGCAAGCGTGACAC
CAAGTGCCAGCCCCAGGGCTCCTCAGGGGAGGAGAAGGGGACCCGACCACACTCAGGGGTGGAGAAGCATCTGA
AAGAAAAAGGCCGACTCGGGCTGTTCAACTTCAAAGACACCAAGTACCAGTCCGGTGTACGTCAATCCGAGGA
GAAGGATGAGTGCGTCATAGCAACTGAGGTGTAAATGGAAGTGAGATGGCAAGACTCCCGTTTCTCTTAAATA
AGTAAATTTCAAGGATATATGCCCCAACGAATGCTGCTGAAGAGGAGGAGGCCTCGTGGACTGCTGCTGAGAA
ACCGAGTTTACACCGAGCAGGTTCTCCTCCTGAGGTCTCGACGCCTGCCGACAGCCTGTGCGGGCCCCGGCCGCC
TGCGGCACTGCCCTCCGTGACGTGCGCTTGCACTATGGACAGTTGCTCTTAAGAGAATATATATTTAAATGGGT
GAACTGAATTACGCATAAGAAGCATGCACTGCCTGAGTGTATATTTTGGATTCTTATGAGCCAGTCTTTTCTTGA
ATTAGAAACACAAACTGCCTTTATTTGCTCTTTTGTACGAAGATGTGCTTTTTCTAGATGGAAAGATGTGT
GTTATTTTTTGGATTTGTAAATATTTTTTATGATATCTGTAAAGCTTGAGTATTTTGTGATGTTCTGTTTTTA
TAATTTAAATTTTGGTAAATATGTACAAAGGCACCTCGGGTCTATGTGACTATATTTTTTGTATATAAATGTAT
TTATGGAATATTGTGCAATGTTATTTGAGTTTTTTACTGTTTTTGTAAATGAAGAAATTCCTTTTTAAATATTT
TTCCAAAATAAATTTTATGAATGACAAAAA

346/550

FIGURE 346

MGSRCALALAVLSALLCQVWSSGVFELKLQEFVNKKGLLGNRNCCRGAGPPPCACRTFFRVC
LKHYQASVSPEPPCTYGS AVTPVLGVDSFSLPDGGGADSAFSPNPIRFPFGFTWPGTFSLIEA
LHTDSPDDLATENPERLISRLATQRHLTVGEEWSQDLHSSGRDLDKYSYRFVCDEHYGEGCS
VFCRPRDDAFGHFTCGERGEKVCNPGWKGPYCTEPICLPGCDEQHGFCDKPGECKCRVWQGR
YCDECIRYPGCLHGTCQQPWQCNCQEGWGLFCNQDLNYCTHHKPCKN GATCTNTGQGSYTCS
CRPGYTGATCELGIDECDPSPCKNGGSCTDLENSYSCTCPPGFYKICELSAMTCADGPCFNG
GRCSDSPDGGYSCRCPVGYSGFNCEKKIDYCSSSPCSNGAKCVDLGDAYLCRCQAGFSGRHCD
DNVDDCASSPCANGGTCTRDGVNDFSTCTPPGYTGRNCSAPVSRCEHAPCHNGATCHERGHRYV
CECARGYGGPNCQFLLPELPPGPAVVDLTEKLEGQGGPFPWVAVCAGVILVLMLLLGCAAVVV
CVRLRLQKHRPPADPCRGETETMNNLANCQREKDISVSIIGATQIKNTNKKADFGDHSADKN
GFKARYPAVDYNLVQDLKGDDTAVRDAHSKRDTKCQPOGSSGEEKGTPPTLRGGEASERKRPD
SGCSTSKDTKYQSVYVISEEKDECVIATEV

Important features:**Signal sequence:**

Amino acids 1-21

Transmembrane domain:

Amino acids 546-566

N-glycosylation site:

Amino acids 477-481

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 660-664

Tyrosine kinase phosphorylation sites:

Amino acids 176-185;252-261

N-myristoylation sites:

Amino acids 2-8;37-43;40-46;98-104;99-105;262-268;281-287;
282-288;301-307;310-316;328-334;340-344;378-384;387-393;512-518;
676-682;683-689;695-701

Aspartic acid and asparagine hydroxylation sites:

Amino acids 343-355;420-432;458-470

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 552-563

EGF-like domain cysteine pattern signature:

Amino acids 243-255;274-286;314-326;352-364;391-403;429-441;
467-479;505-517

347/550

FIGURE 347

CCCACGCGTCCGCACCTCGGCCCCGGGCTCCGAAGCGGCTCGGGGGCGCCCTTTCGGTCAACA
TCGTAGTCCACCCCCTCCCCATCCCCAGCCCCGGGGATTTCAGGCTCGCCAGCGCCCAGCCAG
GGAGCCGGCCGGGAAGCGCGATGGGGGCCCCAGCCGCCTCGCTCCTGCTCCTGCTCCTGCTGT
TCGCCTGCTGCTGGGCGCCCGGCGGGGCCAACCTCTCCCAGGACGACAGCCAGCCCTGGACAT
CTGATGAAACAGTGGTGGCTGGTGGCACCGTGGTGTCAAGTGCCAAGTGAAAGATCACGAGG
ACTCATCCCTGCAATGGTCTAACCTGCTCAGCAGACTCTCTACTTTGGGGAGAAGAGAGCCC
TTCGAGATAATCGAATTCAGCTGGTTACCTCTACGCCCCACGAGCTCAGCATCAGCATCAGCA
ATGTGGCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCCTGTGCGAACTG
CCAAGTCCCTCGTCACTGTGCTAGGAATTCACAGAAGCCCATCATCACTGGTTATAAATCTT
CATTACGGGAAAAAGACACAGCCACCCTAAACTGTCTAGTCTTCTGGGAGCAAGCCTGCAGCCC
GGCTCACCTGGAGAAAGGGTGACCAAGAACTCCACGGAGAACCAACCCGCATACAGGAAGATC
CCAATGGTAAACCTTCACTGTCAGCAGCTCGGTGACATTCCAGGTTACCCGGGAGGATGATG
GGGCGAGCATCGTGTGCTCTGTGAACCATGAATCTCTAAAGGGAGCTGACAGATCCACCTCTC
AACGCATTGAAGTTTTATACACACCAACTGCGATGATTAGGCCAGACCCTCCCCATCCTCGTG
AGGGCCAGAAGCTGTTGCTACACTGTGAGGGTCGCGGCAATCCAGTCCCCAGCAGTACCTAT
GGGAGAAGGAGGGCAGTGTGCCACCCCTGAAGATGACCCAGGAGAGTGCCCTGATCTTCCCTT
TCCTCAACAAGAGTGACAGTGGCACCTACGGCTGCACAGCCACCAGCAACATGGGCAGCTACA
AGGCCTACTACACCCTCAATGTTAATGACCCCAGTCCGGTGCCCTCCTCCTCCAGCACCTACC
ACGCCATCATCGGTGGGATCGTGGCTTTTCATTGTCTTCTGCTGCTCATCATGCTCATCTTCC
TTGGCCACTACTTGATCCGGCACAAAGGAACCTACCTGACACATGAGGCAAAAGGCTCCGACG
ATGCTCCAGACGCGGACACGGCCATCATCAATGCAGAAGGCGGGCAGTCAGGAGGGGACGACA
AGAAGGAATATTTTCATCTAGAGGCGCCTGCCCACCTTCTGCGCCCCCAGGGGCCCTGTGGGG
ACTGCTGGGGCCGTCACCAACCCGGACTTGTACAGAGCAACCGCAGGGCCGCCCTCCCGCTT
GCTCCCCAGCCCACCCACCCCTGTACAGAATGTCTGCTTTGGGTGCGGTTTTGTACTCGGT
TTGGAATGGGGAGGGAGGAGGGCGGGGGAGGGGAGGGTTGCCCTCAGCCCTTCCGTGGCTT
CTCTGCATTTGGGTATTATTATTTTGTAAACAATCCCAAATCAAATCTGTCTCCAGGCTGGA
GAGGCAGGAGCCCTGGGGTGAGAAAAGCAAAAAACAAACAAAAACA

348/550

FIGURE 348

MGAPAASLLLLLLLLFACCWAPGGANLSQDDSQPWTSDETVVAGGTVVVKCQVKDHEDSSLQWS
NPAQQTLTYFGEKRALRDNRQLVTSTPHELSSISINVALADEGEYTCSTFTMPVRTAKSLVTV
LGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRIQEDPNGKTFT
VSSSVTFQVTTREDDGASIVCSVNHESLKGADRSTSQRIEVLYTPTAMIRPDPPHPREGQKLLL
HCEGRGNPVPQQYLWEKEGSPPLKMTQESALIFPFLNKSDSGTYGCTATSNMGSYKAYYTLN
VNDPSPVPSSSSTYHAIIGGIVAFIVFLLLIMLIIFLGHYLIRHKGYLTAEAKGSDDAPDADT
AIINAEGGQSGGDDKKEYFI

Important features:**Signal sequence:**

amino acids 1-20

Transmembrane domain:

amino acids 331-352

N-glycosylation site.

amino acids 25-29, 290-294

Casein kinase II phosphorylation site.

amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

N-myristoylation site.amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304,
306-310, 334-340, 360-364, 385-389, 386-390**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 7-18

349/550

FIGURE 349

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATCTCCCTGTTGAATTTTTTGCACATGGAGGACAGCAGCAAAG
AGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTACCATACGCCCTCAGGACGTTCCCTCTA
GCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGTCATTTTGATTTTGCTGTTATTTTTTTTTCTTTTTCTT
TTTCCCACCACATTGTATTTTATTTCCGTACTTCAGAAATGGGCCTACAGACCACAAAGTGGCCAGCCATGGGG
CTTTTTTCTGAAGTCTTGGCTTATCATTTCCTGGGGCTCTACTCACAGGTGTCCAACTCCTGGCCTGCCCTA
GTGTGTGCCGCTGCGACAGGAACCTTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTGCCTCTTGGGATCCCGG
AGGGCGTAACCGTACTCTACCTCCACAACAACCAAATTAATAATGCTGGATTTCCTGCAGAACTGCACAATGTAC
AGTCGGTGACACGGTCTACCTGTATGGCAACCACTGGACGAATTCCCATGAACCTTCCCAAGAATGTAGAG
TTCTCCATTTCAGGAAAACAATATTCAGACCATTTCACGGGCTGCTCTTGCCAGCTCTTGAAGCTTGAAGAGC
TGCACCTGGATGACAACTCCATATCCACAGTGGGGGTGGAAGACGGGGCCTTCCGGGAGGCTATTAGCCTCAAAT
TGTGTTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCCTGTGGACTTGCAAGAGCTGAGAGTGG
ATGAAAAATCGAATTGCTGTCATATCCGACATGGCCTTCCAGAATCTCACGAGCTTGAGCGTCTTATTGTGGACG
GGAACCTCCTGACCAACAAGGGTATCGCCGAGGGCACCTTCAGCCATCTCACCAAGCTCAAGGAATTTTCAATTG
TACGTAATTGCTGTCCACCCCTCCTCCCGATCTCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACC
AGATAAACCACATTCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCAACAACCAAC
TGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACTGAAGCAGCTCACTGCTCGGAATAACCCCTGGT
TTTGTGACTGCAGTATTAAATGGGTACAGAATGGCTCAAATATATCCCTTCATCTCTCAACGTGCGGGGTTTCA
TGTGCCAAGGTCTGAACAAGTCCGGGGGATGGCCGTGAGGAATTAATATGAATCTTTGTCTGTCCACCA
CGACCCCGGCCTGCCTCTCTTACCCCGAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCACCCCTCTCTA
TTCCAAACCCTAGCAGAAGCTACAGCCTCCAACCTCTACCACATCGAACTTCCACGATTCTGACTGGGATG
GCAGAGAAAGAGTGACCCACCTATTTCTGAACGGATCCAGCTCTCTATCCATTTTGTGAATGATACTTCCATTC
AAGTCAGCTGGCTCTCTCTCTTACCGTGATGGCATACAAACCTCACATGGGTGAAAATGGGCCACAGTTTAGTAG
GGGGCATCGTTCAGGAGCGCATAGTCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCA
CCTATCGGATTTGTTTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACACCATTTGTTAGAGGCCA
CCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGACGTCCACAGCATGG
GCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATATTTGTGCTGGTGGTCTTGCTCAGCGTCTTTT
GCTGGCATATGCACAAAAGGGGCGCTACACCTCCAGAAAGTGGAAATACAACGGGGCGGGCGGAAAGATGATT
ATTGCGAGGCAGGCACCAAGAAGGACAACCTCCATCCTGGAGATGACAGAAACAGTTTTTCAGATCGTCTCTTAA
ATAACGATCAACTCCTTAAAGGAGATTTAGACTGCAGCCATTTACACCCCAAATGGGGGCATTAATTACACAG
ACTGCCATATCCCCAACACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGCACTGCCATACGTCAGCAGC
CAGAGGCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAACACACTCGTGTGTGCACATAAAGACACGCAG
ATTACATTTGATAAATGTTACACAGATGCATTTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAA
TGGGATTTAAAAAAGTGCTATCTTTCTATTTCAAGTTAATTACAAACAGTTTTGTAACTCTTGCTTTTTAA
TCTT

350/550

FIGURE 350

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIPE
GVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENNIQTISR
AALAQLLKLEELHLDNDSISTVGVEDGAFREAI SLKLLFLSKNHLSSVPVGLPVDLQELRVDE
NRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPDLPGT
HLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNNPWFCD
CSIKWVTEWLKYIPSSLNVRGFMCGQPEQVRGMARELNMNLLSCPTTTPGLPLFTPAPSTAS
PTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWDGRERVTPPISERIQLSIHFVNDTSIQVSW
LSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPLDAFNRYRAV
EDTICSEATTHASYLNNGSNTASSHEQTTSMSGSPFLLAGLIGGAVIFVLVVLVSVFCWHMH
KKGRTYSQKWKNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKGD FRLQPIYTP
NGGINYTDCHIPNNMRYCNSSVPDLEHCHT

Important features:**Signal peptide:**

amino acids 1-42

Transmembrane domain:

amino acids 542-561

N-glycosylation site.

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

Casein kinase II phosphorylation site.

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

Tyrosine kinase phosphorylation site.

amino acids 319-328

N-myristoylation site.amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,
522-528, 545-551, 633-639**Amidation site.**

amino acids 581-585

Leucine zipper pattern.

amino acids 164-186

Phospholipase A2 aspartic acid active site.

amino acids 39-50

351/550

FIGURE 351

AGCCGACGCTGCTCAAGCTGCAACTCTGTTGCAGTTGGCAGTTCTTTTCGGTTTCCCTCCTGCTGTTTGGGGGCA
TGAAAGGGCTTCGCCGCCGGGAGTAAAGAGGAATTGACCGGGCAGCGCGAGGGAGGAGCGCGCACGCGACCGC
GAGGGCGGGCGTGCACCCCTCGGCTGGAAGTTTGTGCCGGGCCCGAGCGCGCGCGGCTGGGAGCTTCGGGTAGA
GACCTAGGCCGCTGGACCGCGATGAGCGCGCCGAGCCTCCGTGCGCGCGCGCGGGGTGGGGCTGCTGCTGTGC
GCGGTGCTGGGGCGCGCTGGCCGGTCCGACAGCGCGGTGCGGGGAACTCGGGCAGCCCTCTGGGGTAGCCGCC
GAGCGCCCATGCCCACTACCTGCCGTGCTCGGGGACCTGCTGGACTGCAGTCGTAAGCGGCTAGCGCGTCTT
CCCGAGCCACTCCCGTCCGTGGGTGCTCGCTCGGCTGGAAGTGAAGTGAACAACTGAATGGAGACCATTCCAAATCTGGGACCA
GTCTCGGCAAATATTACACTTCTCTCCTTGGCTGGAACAGGATTGTTGAAATACTCCCTGAACATCTGAAAGAG
TTTCAGTCCCTTGAACTTTGGACCTTAGCAGCAACAATATTTCAGAGCTCCAACTGCATTTCAGGCCCTACAG
CTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAATTTGGCCAAACACA
CTCCTTGTGTTAAAGCTGAACAGGAACCGAATCTCAGTATCCCACCAAGATGTTTAACTGCCCCAACTGCAA
CATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGGTGCTCTGAAGTCT
CTGAAATGCAAAGAAATGGAGTAACGAACTTATGGATGGAGCTTTTGGGGCTGAGCAACATGGAAATTTTG
CAGCTGGACCATAACAACCTAACAGAGATTACCAAAGGCTGGCTTTACGGCTTGCTGATGCTGCAGGAACCTCAT
CTCAGCCAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAGTGAGCTGGACCTA
ACTTCAATCACTTATCAAGGTTAGATTCAAGCTTCCTTGGCCTAAGCTTACTAAATACACTGCACATTGGG
AACACAGAGTCAGCTACATTGCTGATTGTGCTTCCGGGGCTTCCAGTTTAAAGACTTTGGATCTGAAGAAC
AATGAAATTTCTGGACTATTGAAGACATGAATGGTGCTTCTCTGGGCTTGACAACTGAGGCGACTGATACTC
CAAGGAAATCGGATCCGTTCTATTACTAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCATCTAGACCTGAGT
GACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCAAAATGAAGAACTGCAACAATTCGATTTAAATACA
TCAAGCCTTTTGTGCGATTGCCAGCTAAAATGGCTCCACAGTGGGTGGCGGAAACAACTTTAGAGCTTTGTA
AATGCCAGTTGTGCCCATCTCAGCTGCTAAAGGAGAAGCAATTTTGTCTGTAGCCAGATGGCTTTGTGTGT
GATGATTTTCCCAAACCCAGATCACGGTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTCCAATTTGAGTTTC
ATCTGCTCAGCTGCCAGCAGCAGTATCCCAATGACTTTTGTCTGGAAAAAGACAATGAAGTACTGCATGAT
GCTGAAATGGAAATTTATGCACACCTCCGGGCCCCAAGGTGGCGAGGTGATGGAGTATACCACCATCTTCGGCTG
GCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCATCTCCAATCACTTTGGTTTATCCTACTCTGTG
AAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTCACCAAGACCCCATGGATCTCACCATCCGAGCTGGGGCC
ATGGCAGCTTGGAGTGTGCTGCTGTTGGGGCACCCAGCCCCCAGATAGCCTGGCAGAAGGATGGGGGCACAGAC
TTCCCAGCTGCACGGGAGAGACGCATGCTGATGCCCGAGGATGACGTGTTCTTATCGTGGATGTGAAGATA
GAGGACATTTGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTAGCAAATGCAACTCTGACTGTC
CTAGAAACACCATCATTTTTTGGGCCACTGTTGGACCGAACTGTAACCAAGGGAGAAACAGCCGCTCTACAGTGC
ATTGCTGGAGGAAGCCCTCCCCCTAACTGAAGTGGACCAAGATGATAGCCCATTTGGTGGTAACCGAGAGGCAC
TTTTTTCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGTGATGCTGGGAAATACACATGTGAG
ATGTCTAACACCCCTTGGCACTGAGAGAGGAAACGTGCGCCTCAGTGTGATCCCACTCCAACCTGCGACTCCCT
CAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGCCGTGGTTTGTGT
GTGGTGGGCACGTCACTCGTGTGGGTGGTTCATATACACACAAGGCGGAGGAATGAAGATTGCAGCATTACC
AACACAGATGAGACCAACTTGGCAGCAGATATTCTAGTTATTGTGTCATCTCAGGGAACGTTAGCTGACAGGCAG
GATGGGTACGTGTCTTCAGAAAGTGAAGCCACCACAGTTTGTACATCTTCAGGTGCTGGATTTTCTTACCA
CAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGCCACAGATCTGTTT
CTTTGTCGGTTTTTGGGATCCACAGGCCCTATGATTTGAAGGAAATGTGTATGGCTCAGATCCTTTTGAACA
TATCATACAGGTTGCGTCTGACCCAAGAACAGTTTTAATGGACCACTATGAGCCCAGTTACATAAAGAAAAAG
GAGTGCTACCCATGTTCTCATCTCAGAAGAATCCTGCGAACGGAGCTTCAGTAATATATCGTGGCCTTCACAT
GTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGGACCTGGAATGAAAAATCTGTGTCTAAACAAGTCC
TCTTTAGATTTTGTAGTCAAAATCCAGAGCCAGCGTCCGTTGCTCGAGTAATCTTTTATGGGTACCTTTGGAAAA
GCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTGAGCCAAGAGCCTTTTAT
TTGAAGCTCATTCTTCCCGAGACTTGGACTCTGGGTGAGGGAAGATGGGAAAGAAAGGACAGATTTTCAGGAA
GAAATCAGATTTGTACCTTTAAACAGACTTTAGAAACTACAGGACTCCAAATTTTCACTTATGACTTTGGAC
ACATAGACTGAATGAGACCAAGGAAAGCTTAACATACTACCTCAAGTGAACCTTTATTTAAAGAGAGAGAAT
CTTATGTTTTTTAAATGGAGTTATGAATTTTAAAGGATAAAAAATGCTTTATTTATACAGATGAACCAAAATTAC
AAAAAGTTATGAAAATTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTTAACTATTTTTTAACTTTG
TTTTATGCAAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTATGTATTTTATAATGCCAGA
TTTCTTTTATGAAAATGAGTTACTAAAGCATTTTAAATAAATACCTGCCTTGTAACATTTTAAATAGAACTT
ACTTCATTATATTTTGCACATTATATTTAATAAATGTGTCAATTTGAAAAAAGAAAAAAGAAAAAAGAAAAA

352/550

FIGURE 352

MSAPSLRARAAGLGLLLCAVLGRAGRSDSGGRGELGQPSGVAAERPCPTTCRCCLGDLDCSRKRLARLPEPLPSW
VARLDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSANITLLSLAGNRIVEILPEHLKEFQSLETL
DLSSNNISELQTAFAPALQKYLNSNRVTSMEPGYFDNLANTLLVLKLNRRNRIAPPKMFKLPQLQHLELNRN
KIKNVDGLTFQGLGALKSLKMQRNGVTKLMDGAFWGLSNMEILQLDHNNTLITKGWLYGLLMLQELHLSQNAIN
RISPDWEFCQKLSELDLTFNHL SRLDDSSFLGSLNLT LHIGNNRVSYIADCAFRGLSSLKTLDLKNNEISWTI
EDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDSLDAIMSLQGNAFSQMKKLQQLHLNTSSLLCDC
QLKWL PQWVAENNFQSFVNASCAHPQLLKGRSIFAVSPDGFVCDDFPKPQITVQPETQSAIKGSNLSFICSAASS
SDSPMTFAWKDNEELLHDAEMENYHLRAQGGVEYTTILRLREVEFASEGKYQCVISNHFSSYSVKAKLTVN
MLPSFTKTPMDLTIRAGAMARLECAAVGHPAPQIAWQKDGGTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYS
CTAQSAGSISANATLTVLETPSFLRPLLDRTVTKGETAVLQCIAGGSPPPKNLWTKDDSPLVVTERHFFAAGNQ
LLIIVDSVDSDAGKYTCEMSNTLGTGRGNVRLSVIPTCTCDSPQMTAPSLDDDGWATVGVVIIAVVCCVVGTSLV
WVVIYHTRRRNEDCSITNTDETNPADIPSYLSSQGTADRDQGYVSSSESGSHHQFVTSSGAGFFLPQHDSST
CHIDNSSEADVEAATDLFLCPFLGSTGPMYLGKNVYGSDPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCSH
PSEESCERSFSNISWPSHVRKLNTSYSHNEGPGMKNLCLNKSSLD SANPEPASVASSNSFMGTFGKALRRPHL
DAYSSFGQSDCQPRAFYLKAHSSPDLDGSGSEEDGKERTDFQEEHNICTFKQTLNRYTNPNFQSYDLDT

Important features:**Signal sequence:**

amino acids 1-27

Transmembrane domain:

amino acids 808-828

N-glycosylation site.amino acids 122-126, 156-160, 274-278, 442-446, 469-473, 515-519,
688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020**Glycosaminoglycan attachment site.**

amino acids 886-890

Casein kinase II phosphorylation site.amino acids 99-103, 180-184, 263-267, 314-318, 324-328, 374-378,
383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735,
799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022,
1073-1077, 1079-1083, 1081-1085**Tyrosine kinase phosphorylation site.**

amino acids 667-675

N-myristoylation site.amino acids 14-20, 36-42, 239-245, 257-263, 380-386, 427-433,
513-519, 588-594, 672-678, 683-687, 774-780, 933-939**Leucine zipper pattern.**

amino acids 58-80, 65-87

353/550

FIGURE 353

GGGGGTTAGGGAGGAAGGAATCCACCCCCACCCCCCAAACCCTTTTCTTCTCCTTTCTGGCTTCGGACATTGG
AGCACTAAATGAACCTGAATTGTGTCTGTGGCGAGCAGGATGGTCGCTGTTACTTTGTGATGAGATCGGGGATGA
ATTGCTCGCTTTAAAAATGCTGCTTTGGATTCTGTTGCTGGAGACGTCTCTTTGTTTTGCCGCTGGAAACGTTAC
AGGGGACGTTTGCAAAGAGAAGATCTGTTCTGCAATGAGATAGAAGGGGACCTACACGTAGACTGTGAAAAAA
GGGCTTCACAAGTCTGCAGCGTTTCACTGCCCGACTTCCAGTTTTACCATTATTTCTGCATGGCAATTCCCT
CACTCGACTTTTCCCTAATGAGTTCGCTAACTTTTATAATGCGGTTAGTTTGACATGGAAAACAATGGCTTGCA
TGAAATCGTTCCGGGGGCTTTTCTGGGGCTGCAGCTGGTGAAAAGGCTGCACATCAACAACAACAAGATCAAGTC
TTTTCGAAAGCAGACTTTTCTGGGGCTGGACGATCTGGAATATCTCCAGGCTGATTTTAATTTATTACGAGATAT
AGACCCGGGGGCTTCCAGGACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACAATCTCATCAGCACCCCTACC
TGCCAACGTGTTCCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAACAGGCTGAAAACGCTGCCCTATGA
GGAGGTCTTGGAGCAAATCCCTGGTATTGCGGAGATCTGCTAGAGGATAACCCTTGGGACTGCACCTGTGATCT
GCTCTCCCTGAAAGATGGCTGGAAAACATTCCCAAGATGCCCTGATCGGCCGAGTGGTCTGCGAAGCCCCAC
CAGACTGCAGGGTAAAGACCTCAATGAAACCACCGAACAGGACTTGTGTCTTTGAAAACCGAGTGGATTCTAG
TCTCCCGGCGCCCCCTGCCAAGAAGAGACCTTTGCTCCTGGACCCCTGCCAACTCCTTTCAAGACAAATGGGCA
AGAGGATCATGCCACACCAGGCTGCTCCAAACGGAGGTACAAAGATCCCAGGCAACTGGCAGATCAAAATCAG
ACCCACAGCAGCGATAGCGACGGGTAGCTCCAGGAACAAACCCTTAGCTAACAGTTTACCCTGCCCTGGGGGCTG
CAGCTGCGACCACATCCCAGGTCGGGTTTAAAGATGAACTGCAACAACAGGAACGTGAGCAGCTTGGCTGATTT
GAAGCCCCAAGCTCTCTAACGTGCAGGAGCTTTTCTACGAGATAACAAGATCCACAGCATCCGAAAATCGCACTT
TGTGGATTACAAGAACCTCATTCTGTTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACAACTTTCAA
GAACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGGAGAAATTCGCGGG
GCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCATCTCCCGGGCACTTTCAATGCCAT
GCCCCAACTGAGGATCCTCATTTCTAACAACAACCTGCTGAGGTCCCTGCCTGTGGACGTGTTGCTGGGGTCTC
GCTCTCTAAACTCAGCCTGCACAACAATTACTTCATGTACCTCCCGGTGGCAGGGGTGCTGGACCAGTTAACCTC
CATCATCCAGATAGACCTCCACGGAAACCCCTGGGAGTGCTCCTGCACAATTGTGCCTTTCAAGCAGTGGGCAGA
ACGCTTGGGTTCCGAAGTGCTGATGAGCGACCTCAAGTGTGAGACGCCGGTGAACCTCTTTAGAAAGGATTTTAT
GCTCCTCTCCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCTCGCCACGTTAACTTCGCACAGTAAAAA
CAGCACTGGGTTGGCGGAGACCGGGACGCACTCCAACCTCTACCTAGACACCAGCAGGGTGTCCATCTCGGTGTT
GGTCCCGGGACTGCTGCTGGTGTGTTGTACCTCCGCCTTCACCGTGGTGGCATGCTCGTGTTTATCCTGAGGAA
CCGAAAGCGGTCCAAGAGACGAGATGCCAACTCCTCCGCGTCCGAGATTAATCCCTACAGACAGTCTGTGACTC
TTCCTACTGGCACAATGGGCCTTACAACGAGATGGGGCCACAGAGTGTATGACTGTGGCTCTCACTCGCTCTC
AGACTAAAGACCCCAACCCCAATAGGGGAGGGCAGAGGGAAGGCGATACATCCTTCCCCACCGCAGGCACCCGGG
GGCTGGAGGGGCGTGTACCCAAATCCCCGCGCCATCAGCCTGGATGGGCATAAGTAGATAAATAACTGTGAGCTC
GCACAACCGAAAGGGCCTGACCCCTTACTTAGCTCCCTCCTTGAAACAAAGAGCAGACTGTGGAGAGCTGGGAGA
GCGCAGCCAGCTCGCTCTTTGCTGAGAGCCCTTTTGACAGAAAGCCAGCACGACCCTGCTGGAAGAAGTACA
GTGCCCTCGCCCTCGGCCCCGGGGCCTGTGGGGTTGGATGCCGCGGTTCTATACATATATACATATATCCACATC
TATATAGAGAGATAGATATCTATTTTCCCTGTGGATTAGCCCCGTGATGGCTCCCTGTTGGCTACGCAGGGAT
GGGCAGTTGCACGAAGGCATGAATGTATTGTAAATAAGTAACTTTGACTTCTGAC

354/550

FIGURE 354

MLLWILLLETSLCFAAGNVTGDVCKEKICSCNEIEGDLHVDCEKKGFTSLQRFTAPTSQFYHL
FLHGNSLTRLFPNEFANFYNAVSLHMENGLHEIVPGAFLGLQLVKRLHINNNKIKSFRKQTF
LGLDDLEYLQADFNLLRDIDPGAQDLNKLEVLILNDNLISTLPANVFQYVPITHLDLRGNRL
KTLPEYEEVLEQIPGIAEILLEDNPWDCTCDLLSLKEWLENIPKNALIGRVVCEAPTRLQGKDL
NETTEQDLCPLKNRVDSSLPAPPAQEETFAPGGLPTPFKTNGQEDHATPGSAPNGGTKIPGNW
QIKIRPTAAIATGSSRNKPLANSPLCPGGCSCDHIPGSGLKMNCCNNRVSSSLADLKPKLSNVQ
ELFLRDNKIHSIRKSHFVDYKNLILLDLGNNNIATVENNTFKNLLDLRWLYMDSNYLDTLSRE
KFLAGLQNLLEYLNVEYNAIQLILPGTFNAMPKLRILILNNNLLRSLPVDVFAGVSLSKLSLHNN
YFMYLPVAGVLDQLTSIIQIDLHGNPWECSCTIVPFKQWAERLGSEVLMSDLKCETPVNFFRK
DFMLLSNDEICPQLYARISPTLTSHSKNSTGLAETGTHSNSYLDTSRVSISVLVPGLLLVFVT
SAFTVVGMLVFILNRKRKRKRRDANSSASEINSLQTVCDSSYWHNGPYNADGAHRVYDCGSHS
LSD

Important features:**Signal sequence:**

amino acids 1-15

Transmembrane domain:

amino acids 618-638

N-glycosylation site.

amino acids 18-22, 253-257, 363-367, 416-420, 595-599, 655-659

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 122-126, 646-650

Casein kinase II phosphorylation site.amino acids 30-34, 180-184, 222-226, 256-260, 366-370, 573-577,
608-612, 657-661, 666-670, 693-697**N-myristoylation site.**amino acids 17-23, 67-73, 100-106, 302-308, 328-334, 343-349,
354-360, 465-471, 493-499, 598-604, 603-609**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 337-348

355/550

FIGURE 355

AGTCGACTGCGTCCCCTGTACCCGGCGCCAGCTGTGTTCTGACCCCAAGAATAACTCAGGGCTGCACCGGGCCTG
GCAGCGCTCCGCACACATTTCTGTGCGGGCCTAAGGGAACTGTTGGCCGCTGGGCCCCGGGGGGATTCTTGG
CAGTTGGGGGGTCCGTGCGGAGCGAGGGCGAGGGGAAGGGAGGGGAACCGGGTTGGGGAAGCCAGCTGTAGAG
GGCGGTGACCGCGCTCCAGACACAGCTCTGCGTCTCGAGCGGGACAGATCCAAGTTGGGAGCAGCTCTGCGTGC
GGGGCCTCAGAGAATGAGGCCGGCGTTTCGCCCTGTGCCTCCTCTGGCAGGCGCTCTGGCCCCGGGCCGGGCGCGG
CGAACACCCCACTGCCGACCGTGCTGGCTGCTCGGCCTCGGGGGCCTGCTACAGCCTGCACCACGCTACCATGAA
GCGGCAGGCGGCCGAGGAGGCCTGCATCTGCGAGGTGGGGCGCTCAGCACCGTGCGTGCGGGCGCCGAGCTGCG
CGCTGTGCTCGCGCTCCTGCGGGCAGGCCCAGGGCCCGGAGGGGGCTCCAAAGACCTGCTGTTCTGGGTGCGACT
GGAGCGCAGGCGTTCCCACTGCACCCTGGAGAACGAGCCTTTGCGGGGTTTCTCTGGCTGTCTCCGACCCCGG
CGGTCTCGAAAGCGACACGCTGCAGTGGTGGAGGAGCCCCAACGCTCCTGCACCGCGCGGAGATGCGCGGTACT
CCAGGCCACCGGTGGGGTCGAGCCCGCAGGCTGGAAGGAGATGCGATGCCACCTGCGCGCCAACGGCTACCTGTG
CAAGTACCAGTTTGAGGTCTTGTGTCTGCGCCGCGCCCCGGGGCCGCTCTAACTTGAGCTATCGCGCGCCCTT
CCAGTGCACAGCGCCGCTCTGGACTTCAGTCCACCTGGGACCGAGGTGAGTGCGCTCTGCCGGGGACAGTCCCC
GATCTCAGTTACTTGATCGCGGACGAAATCGGCGCTCGCTGGGACAACTCTCGGGCGATGTGTTGTGTCCCTG
CCCCGGGAGGTACCTCCGTGCTGGCAAATGCGCAGAGTCCCTAACTGCCTAGACGACTTGGGAGGCTTTGCCTG
CGAATGTGCTACGGGCTTCGAGCTGGGGAAGGACGGCCGCTCTTGTGTGACCAGTGGGGAAGGACAGCCGACCCT
TGGGGGGACCGGGTGCCACCAGGCGCCCGCGGCCACTGCAACCAGCCCCGTGCCGCAGAGAACATGGCCAAT
CAGGGTCGACGAGAAGCTGGGAGAGACACCACTTGTCCCTGAACAAGACAATTAGTAACATCTATTCTGAGAT
TCCTCGATGGGGATCACAGAGCACGATGTCTACCCTTCAAATGTCCCTTCAAGCCGAGTCAAAGGCCACTATCAC
CCCATCAGGGAGCGTGATTTCCAAGTTTAATTCTACGACTTCCTCTGCCACTCCTCAGGCTTTCGACTCCTCCTC
TGCCGTGGTCTTCATATTTGTGAGCACAGCAGTAGTAGTGTGGTGATCTTGACCATGACAGTACTGGGGCTTGT
CAAGCTCTGCTTTCAGGAAAGCCCTCTTCCCAGCCAAGGAAGGAGTCTATGGGCCCGCGGGCCTGGAGAGTGA
TCCTGAGCCCGCTGCTTTGGGCTCCAGTTCTGCACATTGCACAAACAATGGGGTGAAAGTCGGGGACTGTGATCT
GCGGGACAGAGCAGAGGTGCCTTGCTGGCGGAGTCCCCTCTGGCTCTAGTGATGCATAGGGAAACAGGGGACA
TGGGCACTCCTGTGAACAGTTTTTCACTTTTGATGAAACGGGGAACCAAGAGGAACCTTACTTGTGTAACGACAA
TTTCTGCAGAAATCCCCCTTCCCTCTAAATTCCCTTTACTCCACTGAGGAGCTAAATCAGAACTGCACACTCCTTC
CCTGATGATAGAGGAAGTGGAAGTGCCTTTAGGATGGTGATACTGGGGGACCGGTAGTGCTGGGGAGAGATATT
TTCTTATGTTTATTTCGAGAAATTTGGAGAAGTGATTGAACTTTTCAAGACATTGGAAACAAATAGAACACAATAT
AATTTACATTAAAAAATAATTTCTACCAAATGGAAAGGAAATGTTCTATGTTGTTTCAGGCTAGGAGTATATTGG
TTCGAAATCCCAGGGAAAAAATAAAAAATAAAAAATTAAAGGATTGTTGAT

356/550

FIGURE 356

MRPAFALCLLWQALWPGPGGGGEHPTADRAGCSASGACYSLHHATMKRQAAEEACILRGGALST
VRAGAE LRAVLALLRAGPGPGGGSKDLLFWVALERRRSHCTLENEPLRGFSWLSSDPGGLESD
TLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLC PAPRPGAASN
LSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGARWDKLSGDVLCPCPGRYLR
AGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTSGEGQPTLGGTGVPTRRPPATATSPVP
QRTWPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITPSGSGVIS
KFNSTTSSATPQAFDSSSAVVFI FVSTAVVVLVILTMTVLGLVKLCFHESPSSQPRKESMGPP
GLESDPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALLAESPLGSSDA

Important features:**Signal sequence:**

amino acids 1-16.

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 189-193, 381-385

Glycosaminoglycan attachment site.

amino acids 289-293

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 98-102, 434-438

Casein kinase II phosphorylation site.

amino acids 275-279, 288-292, 342-346, 445-449

N-myristoylation site.amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157,
185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469,
477-483**Aspartic acid and asparagine hydroxylation site.**

amino acids 262-274

357/550

FIGURE 357

CCCATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGACCTCTACATTCCATTTTGGGAAGA
AGACTAAAAATGCTGTTTCCAATGTGGACACTGAAGAGACAAATTCCTATCCTTTTAAACATAATCCTAATTTCC
AAACTCCTTGGGGCTAGATGGTTTCCCTAAAACTCTGCCCTGTGATGTCACTCTGGATGTTCCAAAGAACCATGTG
ATCGTGGACTGCACAGACAAGCATTGACAGAAATTCCTGGAGGTATTCCCACGAACACCACGAACCTCACCTC
ACCATTAACCACATACCAGACATCTCCCCAGCGTCTTTCACAGACTGGACCATCTGGTAGAGATCGATTTTCAGA
TGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCAGATTAAACCCAGAAGC
TTTAGTGGACTCACTTATTTAAATCCCTTTACCTGGATGGAACCAGCTACTAGAGATACCGCAGGGCCCTCCCG
CCTAGCTTACAGCTTCTCAGCCTTGAGGCCAACACATCTTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCC
AACATAGAAATACTCTACCTGGGCCAAAACCTGTTATTATCGAAATCCTTGTTATGTTTCATATTCAATAGAGAAA
GATGCCTTCTAAACTTGACAAAGTTAAAAGTGCTCTCCCTGAAAGATAACAATGTCACAGCCGTCCCTACTGTT
TTGCCATCTACTTTAACAGAACTATATCTCTACAACAACATGATTGCAAAAATCCAAGAAGATGATTTTAAATAAC
CTCAACCAATTACAAATTCCTGACCTAAGTGGAAATTGCCCTCGTTGTTATAATGCCCCATTTCTTGTGCGCCG
TGTAATAATAATTCTCCCCTACAGATCCCTGTAAATGCTTTTGATGCGCTGACAGAATTAAAAGTTTACCTCTA
CACAGTAACCTCTCTCAGCATGTGCCCCAAGATGGTTTAAAGACATCAACAACTCCAGGAAGTGGATCTGTCC
CAAACTTCTTGGCCAAAGAAATGGGGATGCTAAATTTCTGCATTTTCTCCCCAGCCTCATCCAATTGGATCTG
TCTTTCAATTTTGAAGTTTCTGCTATCGTGCATCTATGAATCTATCACAAGCATTCTTCTCACTGAAAAGCCTG
AAAATTCGCGGATCAGAGGATATGTCTTTAAAGAGTTGAAAAGCTTTAACCTCTCGCCATTACATAATCTTCAA
AATCTTGAAGTTCTTGATCTTGGCACTAAGTTTATAAAAAATTGCTAACCTCAGCATGTTTAAACAATTTAAAGA
CTGAAAGTCATAGATCTTTCAGTGAATAAAATATCACCTTCAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAAAT
GCCAGAACTCTGTAGAAAGTTATGAACCCAGGTCTTGAACAATTACATTATTTAGATATGATAAGTATGCA
AGGAGTTGCAGATTCAAAAACAAAGAGGCTCTTTCATGTCTGTTAATGAAAGCTGCTACAAGTATGGGCAGACC
TTGGATCTAAGTAAAAATAGTATATTTTTTGTCAAGTCTCTGATTTTCAGCATCTTCTTCTCTCAATGCCTG
AATCTGTGAGGAAATCTCATTAGCCAACTCTTAATGGCAGTGAATTCCAACCTTAGCAGAGCTGAGATATTTG
GACTTCTCCAACAACCGGCTTGATTTACTCCATTCAACAGCATTGGAAGAGCTTCACAACTGGAAGTTCTGGAT
ATAAGCAGTAATAGCCATTATTTTCAATCAGAAGGAATTACTCATATGCTAACTTTACCAAGAACCTAAAGGTT
CTGCAGAACTGATGATGAACGACAATGACATCTCTTCTCCACCAGCAGGACCATGGAGAGTGAGTCTCTTAGA
ACTCTGGAATTGAGAGGAATCACTTAGATGTTTTATGGAGAGAAGGTGATAACAGATACTTACAATTATTCAAG
AATCTGCTAAAATTAGAGGAATTAGACATCTCTAAAAATTCCTAAGTTTCTTGCCTTCTGGAGTTTTTGATGGT
ATGCCCTCAAATCTAAAGAATCTCTCTTGGCCAAAAATGGGCTCAAATCTTTCAGTTGGAAGAACTCCAGTGT
CTAAAGAACCTGGAACTTTGGACCTCAGCCACAACCAACTGACCACTGTCCCTGAGAGATTATCCAACCTGTTCC
AGAAGCCTCAAGAATCTGATTCTTAAGAATAATCAAATCAGGAGTCTGACGAAGTATTTTCTACAAGATGCCTTC
CAGTTGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCAGCTTCCCAGAAAATGTCCTC
AACAACTCTGAAGATGTTGCTTTTGCATCATAATCGGTTTTCTGTGCACCTGTGATGCTGTGTGGTTTTGCTGGTGG
GTTAACCATACGGAGGTGACTATTCTTACCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCACACAAGGGC
CAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTCACTTTCCATA
TCTGTATCTCTCTTCTCATGGTGTGATGACAGCAAGTCACTCTATTTCTGGGATGTGTGGTATATTTACCAT
TTCTGTAAGGCCAAGATAAAGGGGTATCAGCGTCTAATATCACCAGACTGTTGCTATGATGCTTTTATTGTGTAT
GACACTAAAGACCCAGCTGTGACCGAGTGGGTTTTGGCTGAGCTGGTGGCCAACTGGAAGACCCAGAGAGAAA
CATTTTAATTTATGTCTCGAGGAAAGGGACTGGTTACCAGGGCAGCCAGTTCTGGAAAACCTTTCCAGAGCATA
CAGCTTAGCAAAAAGACAGTGTGTTGTGATGACAGACAAGTATGCAAAGACTGAAAATTTTAAAGATAGCATTTTAC
TTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGTGATTATCTTGATATTTCTTGAGAAGCCCTTTCAAGAG
TCCAAGTTCTCCAGCTCCGAAAAGGCTCTGTGGGAGTTCTGTCTTGAGTGGCCAAACACCCGCAAGCTCAC
CCATACTTCTGGCAGTGTCTAAAGAACGCCCTGGCCACAGACAATCATGTGGCCTATAGTCAGGTGTTCAAGGAA
ACGGTCTAGCCCTTCTTTGCAAAACACAACCTGCCTAGTTTACCAAGGAGAGGCTGGC

358/550

FIGURE 358

MVFPMTLKRQILILFNIIILISKLLGARWFPKTLPCDVTLDVPKNHVIVDCTDKHLTEIPGGI
PTNTTNLTLTINHIPDISPASFHRLDHLVEIDFRNCNCVPIPLGSKNNMCIKRLQIKPRSFSGL
TYLKSLYLDGNQLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLGQNCYRNP
YVSYSIEKDAFLNLTCLKVLSLKDNNTAVPTVLPSTLTELYLYNNMIAKIQEDDFNNLNQLO
ILDLSGNCPRCYNAPFPCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPPRWFKNINK
LQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNELQVYRASMNLSQAFSSLSLKLIRIR
GYVFKELKSFNLSPLHNLQNLVLDLGTNFIKIANLSMFKQFKRLKVIDLSVNKISPSGDSSE
VGFCSNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQTLDSLKNS
IFFVKSSDFQHLNFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNRDLHLHSTAFEELHK
LEVLDISSNSHYFQSEGITHMLNFTKNLKVQLKMMNDNDISSSTSRTMESESLRTLEFRGNH
LDVLWREGDNRYLQLFKNLLKLEELDISKNLSLPSGVFDGMPPNLKNLSLAKNGLKSFSWK
KLQCLKNLETLDLSHNQLTTVPERLSNCSRSLSKNLILKNNQIRSLTKYFLQDAFQLRYLDLSS
NKIQMIQKTSFPENVLNNLKMLLLHNRFLCTCDVWFVWVWNHTEVTIPYLATDVTGPGA
HKGQSVISLDLYTCELDLTNLILFSLISVSLFLMVMMTASHLYFWDVWYIYHFCKAKIKGYQ
RLISPDCCYDAFIVYDTKDPVTEWVLAELVAKLEDPREKHFNLCEERDWLPGQPVLENLSQ
SIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVILIFLEKPFQKSKFLQLRKRLC
GSSVLEWPTNPQAHFYFWQCLKNALATDNHVAYSQVFKETV

Important features:**Signal sequence:**

amino acids 1-26

Transmembrane domain:

amino acids 840-860

359/550

FIGURE 359

GACGGCTGGCCACCATGCACGGCTCCTGCAGTTTCCTGATGCTTCTGCTGCCGCTACTGCTAC
TGCTGGTGGCCACCACAGGCCCCGTTGGAGCCCTCACAGATGAGGAGAAACGTTTGATGGTGG
AGCTGCACAACCTCTACCGGGCCCAGGTATCCCCGACGGCCTCAGACATGCTGCACATGAGAT
GGGACGAGGAGCTGGCCGCCTTCGCCAAGGCCTACGCACGGCAGTGCGTGTGGGGCCACAACA
AGGAGCGCGGGCGCCGCGGCGAGAATCTGTTGCCATCACAGACGAGGGCATGGACGTGCCGC
TGGCCATGGAGGAGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCTGCAGCC
CAGGCCAGATGTGCGGCCACTACACGCAGGTGGTATGGGCCAAGACAGAGAGGATCGGCTGTG
GTTCCCACTTCTGTGAGAAGCTCCAGGGTGTGAGGAGACCAACATCGAATTACTGGTGTGCA
ACTATGAGCCTCCGGGGAACGTGAAGGGGAAACGGCCCTACCAGGAGGGGACTCCGTGCTCCC
AATGTCCCTCTGGCTACCACTGCAAGAACTCCCTCTGTGAACCCATCGGAAGCCCGGAAGATG
CTCAGGATTTGCCTTACCTGGTAAGTGAAGGCCCATCCTTCCGGGCGACTGAAGCATCAGACT
CTAGGAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATTCCGGCTTTCTTGGTAACAGAGG
TCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGCTGTGGAACCCAGGCCCCAACTTCCTTAG
CAACGAAAGACCCGCCCTCCATGGCAACAGAGGCTCCACCTTGCCTAACAACTGAGGTCCCTT
CCATTTTGGCAGCTCACAGCCTGCCCTCCTTGGATGAGGAGCCAGTTACCTTCCCCAAATCGA
CCCATGTTCTATCCCAAATCAGCAGACAAAGTGACAGACAAAACAAAAGTGCCCTCTAGGA
GCCCAGAGAACTCTCTGGACCCCCAAGATGTCCCTGACAGGGGCAAGGGAACCTCCTACCCCATG
CCCAGGAGGAGGCTGAGGCTGAGGCTGAGTTGCCTCCTTCCAGTGAGGTCTTGGCCTCAGTTT
TTCCAGCCCAGGACAAGCCAGGTGAGCTGCAGGCCACACTGGACCACACGGGGCACACCTCCT
CCAAGTCCCTGCCCAATTTCCCCAATACCTCTGCCACCGCTAATGCCACGGGTGGGCGTGCCC
TGGCTCTGCAGTCGTCCTTGCCAGGTGCAGAGGGCCCTGACAAGCCTAGCGTTGTGTCAGGGC
TGAAGTCCGGCCCTGGTCATGTGTGGGGCCCTCTCCTGGGACTACTGCTCCTGCCTCCTCTGG
TGTTGGCTGGAATCTTCTTGAATGGGATACCACTCAAAGGGTGAAGAGGTGAGCTGTCCTCCTG
TCATCTTCCCCACCCTGTCCCCAGCCCCTAAACAAGATACTTCTTGGTTAAGGCCCTCCGGAA
GGGAAAGGCTACGGGGCATGTGCCTCATCACACCATCCATCCTGGAGGCACAAGGCCTGGCTG
GCTGCGAGCTCAGGAGGCCGCTGAGGACTGCACACGGGGCCACACCTCTCCTGCCCCCTCCC
TCCTGAGTCCTGGGGGTGGGAGGATTTGAGGGAGCTCACTGCCTACCTGGCCTGGGGCTGTCT
GCCACACAGCATGTGCGCTCTCCCTGAGTGCCTGTGTAGCTGGGGATGGGGATTCTTAGGGG
CAGATGAAGGACAAGCCCCACTGGAGTGGGGTTCTTTGAGTGGGGGAGGCAGGGACGAGGGAA
GGAAAGTAACTCCTGACTCTCCAATAAAAACCTGTCCAACCTGTGAAA

360/550

FIGURE 360

MHGSCSFMLLLPLLLLLLVATTGPVGALTDEEKRLMVELHNLRYAQVSPTASDMLHMRWDEEL
AAFAKAYARQCVWGHNKERGRGENLFAITDEGMDVPLAMEEWHHEREHYNLSAATCSPGQMC
GHYTQVWAKTERIGCGSHFCEKLQGVETNIELLVVCNYEPPGNVKGKRPYQEGTPCSQCPSG
YHCKNSLCEPIGSPEDAQDLPYLVTEAPSFRAEASDSRKMGTTPSSLATGIPAFLVTEVSGSL
ATKALPAVETQAPTSLATKDPPSMATEAPPCVTTEVPSILAAHSLPSLDEEPVTFPKSTHVPI
PKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEELPPSSEVLASVFPAQD
KPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGPDKPSVVSGLNNGP
GHVWGPLLGLLLLPLVLGIF

Important features:**Signal sequence:**

amino acids 1-22

N-glycosylation site.

amino acids 114-118, 403-407, 409-413

Glycosaminoglycan attachment site.

amino acids 439-443

Casein kinase II phosphorylation site.

amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

N-myristoylation site.

amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237, 250-256

Amidation site.

amino acids 82-86, 172-176

Peroxidases proximal heme-ligand signature.

amino acids 287-298

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.

amino acids 127-138

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.

amino acids 160-172

361/550

FIGURE 361

GACTAGTTCTCTTGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGACTGG
GGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAACT
GGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGACTGCG
AGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCT**ATG**AGGCCACTCCTCGTCCTGCTGCT
CCTGGGCCTGGCGGCCGGCTCGCCCCACTGGACGACAACAAGATCCCCAGCCTCTGCCCCGGG
GCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGGGGCCGCGATGG
CCGCGACGGCCGCGACGGCGCGCCCGGGGCTCCGGGAGAGAAAGGCGAGGGCGGGAGGCCGGG
ACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCCGCGGGGGCCACCGG
GCCTGCCGGGGAGTGCTCGGTGCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTCCGAGAGCCG
GGTGCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACGAGCAGGGACA
TTACGACGCCGTACCCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTACTTCGCCGTCCA
TGCCACCGTCTACCGGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGAATCCATTGCCTC
TTTCTTCCAGTTTTTTCGGGGGGTGGCCCCAAGCCAGCCTCGCTCTCGGGGGGGCCATGGTGAG
GCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTACATTGGCATCTATGC
CAGCATCAAGACAGACAGCACCTTCTCCGATTTCTGGTGTACTCCGACTGGCACAGCTCCCC
AGTCTTTGCT**TAG**TGCCCACTGCAAAGTGAGCTCATGCTCTCACTCCTAGAAGGAGGGTGTGA
GGCTGACAACAGGTATCCAGGAGGGCTGGCCCCCTGGAATATTGTGAATGACTAGGGAGG
TGGGGTAGAGCACTCTCCGTCCTGCTGCTGGCAAGGAATGGGAACAGTGGCTGTCTGCGATCA
GGTCTGGCAGCATGGGGCAGTGGCTGGATTTCTGCCCAAGACCAGAGGAGTGTGCTGTGCTGG
CAAGTGTAAGTCCCCCAGTTGCTCTGGTCCAGGAGCCACGGTGGGGTGCTCTCTTCCTGGTC
CTCTGCTTCTCTGGATCCTCCCCACCCCTCCTGCTCCTGGGGCCGGCCCTTTTCTCAGAGAT
CACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

362/550

FIGURE 362

MRPLLVLALLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAPG
EKGEGRPGPLPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSKRSESRVPPPSDAPLPFD
RVLVNEQGHYDAVTGKFTCQVPGVYFVAVHATVYRASLQFDLVKNGESIÀSFFQFFGGWPKPA
SLSGGAMVRLEPEDQVWVQVGVDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA

Important features:

Signal sequence.

amino acids 1-15

N-myristoylation sites.

amino acids 11-17, 68-74, 216-222

Cell attachment sequence.

amino acids 77-80

363/550

FIGURE 363

[illegible]

364/550

FIGURE 364

MMWRPSVLLLLLLLLLRHGAQGKPSPDAGPHGQGRVHQAAPLSDAPHDDAHGNFQYDHEAFLGRE
VAKEFDQLTPEESQARLGRIVDRMDRAGDGDGWVSLAELRAWIAHTQQRHIRDSVSAAWDTYD
TDRDGRVGWEELRNATYGHYAPGEEFHDVEDAETYKKMLARDERRFRVADQDGDSMATREELT
AFLHPPEEFPHMRDIVIAETLEDLDRNKDGYVQVEEYIADLYSAEPGEEEPAWVQTERQQFRDF
RDLNKDGHLDGSEVGHWVLPPAQDQPLVEANHLLHESDTDKDGRLSKAEILGNWNMFVGSQAT
NYGEDLTRHHDEL

Important features:**Signal sequence:**

amino acids 1-20

N-glycosylation site.

amino acids 140-144

Casein kinase II phosphorylation site.amino acids 72-76, 98-102, 127-131, 184-188, 208-212, 289-293,
291-295, 298-302**N-myristoylation site.**

amino acids 263-269, 311-317

Endoplasmic reticulum targeting sequence.

amino acids 325-330

365/550

FIGURE 365

GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTCACTGGCCTGATCGCGATGGGGACAAAG
GCGCAAGTCGAGAGGAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGGCA
TTGGGCAGTGTTACAGTGCCTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCTGTG
AAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTGGAAGTTTGACCAAGGA
GACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGGTGACC
TTCTTGCCAACTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATACACTTGT
ATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGTGCTTGTG
CCTCCATCCAAGCCTACAGTTAACATCCCCTCCTCTGCCACCATTGGGAACCGGGCAGTGCTG
ACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGATGGGATAGTG
ATGCCTACGAATCCCAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCCTGAATCCCACA
ACAGGAGAGCTGGTCTTTGATCCCCTGTCAGCCTCTGATACTGGAGAATACAGCTGTGAGGCA
CGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCTGTGGAGCGGAAT
GTGGGGGTCATCGTGGCAGCCGTCCTTGTAACCCTGATTCTCCTGGGAATCTTGGTTTTTGGC
ATCTGGTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGACTTCGAGTAAGAAG
GTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAGACCTCGTCATTCCCTG
GTG**TGAG**CCCTGGTCGGCTCACCGCCTATCATCTGCATTTGCCTTACTCAGGTGCTACCGGACT
CTGGCCCCCTGATGTCTGTAGTTTACAGGATGCCTTATTTGTCTTCTACACCCACAGGGCCC
CCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCCATCCTCCTTCATGCCCTC
CCTCCCTTTCCTACCACTGCTGAGTGGCCTGGAACCTGTTTTAAAGTGTTTATTCCCCATTTCT
TTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCCTTCTAAGTAGACAGCAAAA
TGGCGGGGGTTCGAGGAATCTGCACTCAACTGCCACCTGGCTGGCAGGGATCTTTGAATAGG
TATCTTGAGCTTGGTTCTGGGCTCTTTCCTTGTGTACTGACGACCAGGGCCAGCTGTTCTAGA
GCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGGTGATGACACTGGGGTCCTTCCAT
CTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTGCCACTGGGATCCCTCTGCCCTGTC
CTCCTGAATACAAGCTGACTGACATTGACTGTGTCTGTGGAAAATGGGAGCTCTTGTTGTGGA
GAGCATAGTAAATTTTCAGAGAACTTGAAGCCAAAAGGATTTAAAACCGCTGCTCTAAAGAAA
AGAAAACCTGGAGGCTGGGCGCAGTGGCTCACGCCTGTAATCCCAGAGGCTGAGGCAGGCGGAT
CACCTGAGGTGGGAGTTTCGGGATCAGCCTGACCAACATGGAGAAACCCTACTGGAAATACAA
AGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTCCCAGCTGCTCAGGAGCCTGGCAACAAGAG
CAAAACTCCAGCTCAAAAAAAAAAAAAAAAAA

366/550

FIGURE 366

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSLAYSGFSSPRVEWK
FDQGDTTTRLVCYNNKITASYEDRVTFLLPTGITFKSVTREDTGTYTCMVSEEGGNSYGEVKVKL
IVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPSEYTWFKDGIVMPTNPKSTRAFSNSSYV
LNPTTGELVFDPLSASDTGEYSCEARNGYGTGMTSNAVRMEAVERNVGVIVAAVLVTLILLGI
LVFGIWFAYSRRGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

Important features:**Signal sequence:**

amino acids 1-27

Transmembrane domain:

amino acids 238-255

N-glycosylation site.

amino acids 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 270-274

Casein kinase II phosphorylation site.amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158,
193-197, 203-207, 287-291**N-myristoylation site.**

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

367/550

FIGURE 367

GGGGAGAGGAATTGACCATGTAAAAGGAGACTTTTTTTTTTGGTGGTGGTGGCTGTTGGGTGCCTTGCAAAAATG
AAGGATGCAGGACGCAGCTTCTCCTGGAACCGAACGCAATGGATAAACTGATGTGCAAGAGAGAAGGAAGAAC
GAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATGAAA
TAAACCAGAGTTAGACCCGCGGGGGTGGTGTGTTCTGACATAAAATAAATAATCTTAAAGCAGCTGTTCCCTCC
CCACCCCCAAAAAAGGATGATTGGAATGAAGAACCAGGATTACAAAGAAAAAGTATGTTCAATTTTTCTC
TATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAAAGAACT
GGTGTGGTGGTGTTCCTTTCTTTTGAATTTCCACAAGAGGAGAGGAAATTAATAATACATCTGCAAAAGAAA
TTTCAGAGAAGAAAAGTTGACCGCGGCAGATTGAGGCATTGATTGGGGGAGAGAAAACCAGCAGAGCACAGTTGGA
TTTGTGCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTTCTTCATCAACCTCCTTTTTTTAAAT
TTTTATTCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTCTTAACCACCTGGATTTCCATCTGGATGTTGCT
GTGATCAGTCTGAAATACAACCTGTTTGAATTCAGAAGGACCAACACCAGATAAATTATGAATGTGAACAAGAT
GACCTTACATCCACAGCAGATAATGATAGGTCTAGGTTTAACAGGGCCCTATTGACCCCTGCTTGTGGTGCT
GCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCGGGCTCAGACCTGCCCTTCTGTGTGCTCCTGCAGCAA
CCAGTTTCAGCAAGGTGATTTGTGTTGCGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCCACCAACACACGGCT
GCTGAACCTCCATGAGAACCAATCCAGATCATCAAAGTGAACAGCTTCAAGCAGTTGAGGCAGTTGGAATCCT
ACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTTGGGGCTTTCAATGGTCTGGCGAACCTCAACACTCTGGA
ACTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTGTATACTTGTCTAACTGAAGGAGCTCTGGTT
GCGAAACAACCCCATTTGAAAGCATCCCTTCTTATGCTTTTAAACAGAATTCCCTTCTTGGCGCGACTAGACTTAGG
GGAATTGAAAAGACTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACCTGAGGTATTTGAACCTTGC
CATGTGAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAAACCTAGATGAGCTGGATCTTTCTGGGAATCA
TTTATCTGCCATCAGGCCTGGCTCTTTCCAGGTTTGTATGCACCTTCAAAAACCTGTGGATGATACAGTCCCAGAT
TCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCAACCTGGCACACAATAATCTAAC
ATTACTGCCTCATGACCTCTTCACTCCCTTGATCATCTAGAGCGGATACATTTACATCACAACCTTGGAACTG
TAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTCGAACACAGCTTGTGTGCCCCGGTG
TAACACTCCTCCCAATCTAAAGGGAGGTACATTGGAGAGCTCGACCAGAATTACTTCACATGCTATGCTCCGGT
GATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCTGAGCTGAAATGTGGGGCTCCACATC
CCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACATGGGGCGTACAAAGTGCAGATAGCTGT
GCTCAGTGATGGTACGTTAAATTTACAAATGTAAGTGTGCAAGATACAGGCATGTACACATGTATGGTGAATAA
TTCCGTTGGGAATACTACTGCTTCAGCCACCTGAATGTTACTGCAGCAACCACTACTCCTTTCTTACTTTTC
AACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACGGACCACAGATAACAATGTGGGTCCCACTCC
AGTGGTGCAGTGGGAGACCACCAATGTGACCACCTCTCTCACACCACAGAGCACAAAGGTGACAGAGAAAACCTT
CACCATCCCAGTGACTGATATAAACAGTGGGATCCCAGGAATTGATGAGGTGATGAAGACTACCAAAATCATCAT
TGGGTGTTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCTGGTCATTTTCTACAAGATGAGGAAGCAGACCA
TCGGCAAAACCATCACGCCCCAACAAAGGACTGTTGAAATTATTAATGTGGATGATGAGATTACGGGAGACACACC
CATGGAAGCCACCTGCCCATGCCTGCTATCGAGCATGAGCACCTAAATCACTATAACTCATACAAATCTCCCTT
CAACCACACAACAACAGTTAACACAATAAATCAATACACAGTTCAGTGCATGAACCGTTATTGATCCGAATGAA
CTCTAAAGACAATGTACAAGAGACTCAAATCTAAACATTTTACAGAGTTACAAAAACAAACAATCAAAAAAAA
GACAGTTTATTAATAATGACACAAATGACTGGGCTAAATCTACTGTTTCAAAAAGTGTCTTTACAAAAAACAA
AAAAGAAAAGAAATTTATTTATTAATAATCTATTGTGATCTAAAGCAGACAAAAA

368/550

FIGURE 368

MLNKMTLHPQQIMIGPRFNRAFDPLLVLALLQLLVAGLVRAQTCPSVCSCSNQFSKVICVRKNLREVPDGIS
TNTRLLNLHENQIQIIVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLANLNTLELFDNRLTTIPNGAFVYLSKL
KELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLSYISEGAFEGLSNLRYNLAMCNLREIPNLTPLIKDELD
LSGNHLSAIRPGSFQGLMHLQKLWMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLLH
NPWNCNCDILWLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAELKC
RASTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVNSVGNNTTASATLNVTAATTP
FSYFSTVTVETMEPSQDEARTDNNVGPTPVVDWETTNTTSLTPQSTRSTEKFTTIPVTDINSIGIPGIDEVMKT
TKIIGCFVAITLMAAVMLVIFYKMRKQHRQNHAPTRTVEIINVDEITGDTPMESHLPMPAIEHEHLNHYS
YKSPFNHTTTVTNTINSIHSSVHEPLLIRMNSKDNVQETQI

Important features:**Signal sequence:**

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438, 442-446,
488-492, 606-610**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243, 391-397,
422-428, 433-439, 531-537

369/550

FIGURE 369

CAAACTTGCCTCGCGGAGAGCGCCAGCTTGACTTGAATGGAAGGAGCCCGAGCCCGGAGCGCAGCTGAGAC
TGGGGGAGCGGCTTCGGCCTGTGGGGCGCGCTCGGCGCCGGGGCGCAGCAGGGAAGGGGAAGCTGTGGTCTGCC
CTGCTCCACGAGGCGCACTGGTGTGAACCGGGAGAGCCCTGGGTGGTCCCGTCCCTATCCCTCCTTTATATA
GAAACCTTCCACACTGGGAAGGCAGCGCGAGGCAGGAGGGCTCATGGTGAGCAAGGAGGCCGGCTGATCTGCAG
GCGCACAGCATTCCGAGTTTACAGATTTTACAGATACCAATGGAAGGCGAGGAGGCAGAACAGCCTGCCTGGT
TCCATCAGCCCTGGCGCCAGGCGCATCTGACTCGGCACCCCTGCAGGCACCATGGCCAGAGCCGGGTGCTGC
TGCTCCTGCTGCTGCTGCCGCCACAGCTGCACCTGGGACCTGTGCTTGCCGTGAGGGCCCCAGGATTTGGCCGAA
GTGGCGGCCACAGCCTGAGCCCCGAAGAGAACGAATTTGCGGAGGAGGAGCCGGTGTGGTACTGAGCCCTGAGG
AGCCCGGGCCTGGCCAGCCGCGGTGAGCTGCCCCGAGACTGTGCCTGTTCCAGGAGGGCGCTGCTGGACTGTG
GCGGTATTGACCTGCGTGAGTTCCCGGGGACCTGCCTGAGCACACCAACCACCTATCTCTGCAGAACCAACAGC
TGAAAAGATCTACCTGAGGAGCTTCCCGGCTGCACCGCTGGAGACACTGAACCTGCAAAACACCCGCTGA
CTTCCCGAGGGCTCCAGAGAAGGCGTTTGAGCATCTGACCAACCTCAATTACCTGTACTTTGGCCAATAACAAGC
TGACCTTGGCACCCTGCTTCTGCCAAACGCCCTGATCAGTGTGGACTTTGCTGCCAACTATCTCACCAAGATCT
ATGGGCTCACCTTTGGCCAGAAGCCAACTTGAGGTCTGTGTACCTGCACAAACAAGCTGGCAGAGCCGGG
TGCCGACAAATGTTCAACGGCTCCAGCAACGTCGAGGTCTCATCTGTCCAGCAACTTCTGCGCCACGTGC
CCAAGCACCTGCCGCTGCCCTGTACAAGCTGCACCTCAAGAACAACAAGCTGGAGAAGATCCCCCGGGGGCCT
TCAGCGAGCTGAGCAGCCTGCGCGAGCTATACCTGCAGAACAACTACCTGACTGACGAGGGCCTGGACAACGAGA
CCTTCTGGAAGCTTCCAGCCTGGAGTACCTGGATCTGTCCAGCAACAACCTGTCTCGGGTCCCAGCTGGGCTGC
CGCGCAGCCTGGTGTGCTGCACTTGAGAGAAGAACGCCATCCGGAGCGTGGACGCGAATGTGCTGACCCCATCC
GCAGCCTGGAGTACCTGCTGCTGCACAGCAACCAGCTGCGGGAGCAGGGCATCCACCACTGGCCTTCCAGGGCC
TCAAGCGGTTGCACACGGTGCACCTGTACAACAACGCGTGGAGCGCGTGCCAGTGGCCTGCCTGCGCGCTGC
GCACCCTCATGATCCTGCACAACCAGATCACAGGCATTGGCCGCGAAGACTTTGCCACCACCTACTTCTGGAGG
AGCTCAACCTCAGTACAACCGCATCACAGCCACAGGTGCACCGGACGCTTCCGCAAGCTGCGCCTGCTGC
GCTCGCTGGACCTGTGCGGCAACCGGCTGCACACGCTGCCACCTGGGCTGCCTCGAAATGTCCATGTGCTGAAGG
TCAAGCGCAATGAGCTGGCTGCCTTGGCACGAGGGGCGCTGGCGGGCATGGCTCAGCTGCGTGAGCTGTACCTCA
CCAGCAACCGACTGCGCAGCCGAGCCCTGGGCCCCCGTGCCTGGGTGGACCTCGCCCATCTGCAGCTGCTGGACA
TCGCCGGGAATCAGCTCACAGAGATCCCCGAGGGGCTCCCCGAGTCACTTGAGTACCTGTACCTGCAGAACAA
AGATTAGTGCGGTGCCCCGCAATGCCTTCGACTCCACGCCCAACCTCAAGGGGATCTTTCTCAGGTTTAAACAAGC
TGGCTGTGGGCTCCGTGGTGACAGTGCCTTCCGGAGGCTGAAGCACCTGCAGGTCTTGGACATTGAAGGCAACT
TAGAGTTTGGTGACATTTCCAAGGACCGTGGCCGCTTGGGGAAGGAAAAGGAGGAGGAGGAAGAGGAGGAGGAGG
AGGAAGAGGAAACAAGATAGTGACAAAGGTGATGCAGATGTGACCTAGGATGATGGACCGCCGACTCTTTTCTGC
AGCACACGCCTGTGTGCTGTGAGCCCCCACTCTGCCGTGCTCACACAGACACACCCAGCTGCACACATGAGGCA
TCCCACATGACACGGGCTGACACAGTCTCATATCCCCACCCTTCCCACGGCGTGTCCCACGGCCAGACACATGC
ACACACATCACACCCTCAAACACCCAGCTCAGCCACACACAACCTACCCTCAAACACACACAGTCTCTGTACAC
CCCCACTACCGCTGCCACGCCCTCTGAATCATGCAGGGAAGGGTCTGCCCTGCCCTGGCACACACAGGCACCCA
TTCCCTCCCCCTGCTGACATGTGTATGCGTATGCATACACACCACACACACATGCACAAGTCATGTGCGAA
CAGCCCTCCAAAGCCTATGCCACAGACAGCTCTTGCCCCAGCCAGAATCAGCCATAGCAGCTCGCCGTCTGCCCT
GTCCATCTGTCCGTCCGTTCCCTGGAGAAGACACAAGGGTATCCATGCTCTGTGGCAGGTGCCACCCCTCT
GGAACCTCACAAAGCTGGCTTTTATTCCTTTCCCATCTATGGGGACAGGAGCCTTCAGGACTGCTGGCCCTGGCC
TGGCCACCCCTGCTCCTCCAGGTGCTGGGCGTCACTCTGCTAAGAGTCCCTCCCTGCCACGCCCTGGCAGGACA
CAGGCACTTTTCCAATGGGCAAGCCAGTGGAGGCAGGATGGGAGAGCCCTGGGTGCTGCTGGGGCCTTGGGG
CAGGAGTGAAGCAGAGGTGATGGGGCTGGGCTGAGCCAGGGAGGAAGGACCCAGCTGCACCTAGGAGACACCTT
GTTCTTCAGGCCTGTGGGGGAAGTCCGGGTGCCTTTATTTTATTCTTTCTAAGGAAAAAATGATAAAAAAT
CTCAAAGCTGATTTTCTTGTATAGAAAACTAATATAAAAGCATTATCCCTATCCCTGCAAAAAA

370/550

FIGURE 370

MEGEEAEQPAWFHQWPRPGASDSAPPAGTMAQSRVLLLLLLLLPPQLHLGPNLAVRAPGFGRSG
GHSLSPEENEFAEEEPVLVLSPEEPGPGPAAVSCPRDCACSQEGVVDCCGIDLREFPGDLPEH
TNHLSLQNNQLEKIYPEELSRHRLETLNLQNNRLTSRGLPEKA FEHLTNLNYLYLANNKLT
APRFLPNALISVDFAANYLTKIYGLTFGQKPNLRSVYLHNNKLADAGLPDNMFNGSSNVEVLI
LSSNFLRHVPKHLPPALYKLHLKNNKLEKIPPGAFSELSSLRELYLQNNYLTDEGLDNETFWK
LSSLEYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSVDANVLTPIRSLEYLLLHSNQLREQG
IHPLAFQGLKRLHTVHLYNNALERVPSGLPRRVRTLMLHNQITGIGREDFATTYFLEELNLS
YNRITSPQVHRDAFRKLRLRLSLDLSGNRLHTLPPGLPRNVHVLKVKRNELAALARGALAGMA
QLRELYLTSNRLRSRALGPRAWVDLAHLQLLDIAGNQLTEIPEGLPESLEYLYLQNNKISAVP
ANAFDSTPNLKGIFLRFNKLAVGSVVDSAFRRLKHLQVLDIEGNLEFGDISKDRGLGKEKEE
EEEEEEEEEEETR

Important features:**Signal sequence:**

amino acids 1-48

N-glycosylation site.

amino acids 243-247, 310-314, 328-332, 439-443

Casein kinase II phosphorylation site.

amino acids 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

N-myristoylation site.amino acids 19-25, 107-113, 213-219, 217-223, 236-242, 335-341,
477-483, 498-502, 539-545, 548-554**Leucine zipper pattern.**amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493,
535-557

371/550

FIGURE 371

CACTTTCTCCCTCTCTTCTTTACTTTTCGAGAAACCGCGCTTCCGCTTCTGGTCGCAGAGACCTCGGAGACCGCG
CCGGGAGACGGAGGTGCTGTGGGTGGGGGGACCTGTGGCTGCTCGTACCGCCCCCACCCTCCTTCTTGAC
TGCCGTCTCCGGAAGACCTTTCCCTGCTCTGTTTCTTCCAGGAGTCTGTGCATCGCCCCGACCTGGCCGG
GAGGAGCTTGGCCGGCGGAGATGCTCTAGGGCGCGCGGGAGGAGCGGCCGGCGGGACGGAGGGCCCGGCAG
GAAGATGGGCTCCCGTGGACAGGGAATCTTGTGGCGTACTGCCTGCTCCTTGCCTTTGCCTCTGGCCTGGTCT
GAGTCGTGTGCCCCATGTCCAGGGGGAACAGCAGGAGTGGGAGGGGACTGAGGAGCTGCCGTGCGCTCCGGACCA
TGCCGAGAGGGCTGAAGAACAAACATGAAAAATACAGGCCAGTCAGGACCAGGGGCTCCCTGCTTCCCGGTGCTT
GCGCTGCTGTGACCCCGGTACCTCCATGTACCCGGCGACCGCGTGGCCAGATCAACATCACTATCTTGAAAGG
GGAGAAGGCTGACCGCGGAGATCGAGGCTCCAGGGGAAATATGGCAAAACAGGCTCAGCAGGGGCCAGGGGCCA
CACTGGACCCAAAGGGCAGAAGGGCTCCATGGGGGCCCTGGGGAGCGGTGCAAGAGCCACTACGCCGCTTTTC
GGTGGGCGGAAGAAGCCCATGCACAGCAACCACTACTACCAGACGGTGATCTTCGACACGGAGTTCTGTAACT
CTACGACCACTTCAACATGTTACCGGCAAGTTCTACTGCTACGTGCCCGGCTCTACTTCTTACGCTCAACGT
GCACACCTGGAACCAAGGAGACCTACCTGCACATCATGAAGAACGAGGAGGAGGTGATCTTGTTCGGCA
GGTGGGCGACCGCAGCATCATGCAAGCCAGAGCCTGATGCTGGAGCTGCGAGAGCAGGACCAGGTGTGGGTACG
CCTCTACAAGGGCGAACGTGAGAAGCCATCTTACGCGAGGAGCTGGACACCTACATCACCTTCACTGGCTACCT
GGTCAAGCAGCCACCGAGCCCTAGCTGGCCGGCCACCTCCTTTCTCTCGCCACCTTCCACCCCTGCGCTGTGC
TGACCCACCGCCTCTTCCCGCATCCCTGGACTCCGACTCCCTGGCTTTGGCATTCACTGAGACGCCCTGCACAC
ACAGAAAGCCAAAGCGATCGGTGCTCCAGATCCCGCAGCCTCTGGAGAGAGCTGACGGCAGATGAAATCACCAG
GGCGGGGCACCCGCGAGAACCCTCTGGGACCTTCCGCGGCCCTCTCTGCACACATCCTCAAGTGACCCCGCACGG
CGAGACCGGGGTGGCGGCAGGGCGTCCAGGGTGCGGCACCGCGGCTCCAGTCTTGGAAATAATTAGGCAAATT
CTAAAGTCTCAAAGGAGCAAAGTAAACCGTGAGGACAAAGAAAAGGGTTGTTATTTTGTCTTTCCAGCCAG
CCTGCTGGCTCCCAAGAGAGAGGCCTTTTCACTTGAGACTCTGCTTAAGAGAAGATCCAAAGTTAAAGCTCTGGG
GTCAGGGGAGGGGCGGGGGCAGGAACTACCTCTGGCTTAATTCTTTTAAGCCACGTAGGAACCTTCTTGAGGG
ATAGTGGACCTGACATCCCTGTGGCCTTGCCCAAGGGCTCTGCTGGTCTTTCTGAGTCACAGCTGCGAGGTGA
TGGGGGCTGGGGCCCCAGGCGTCAGCCTCCAGAGGGACAGCTGAGCCCCCTGCCTTGGCTCCAGGTTGGTAGAA
GCAGCCGAAGGGCTCCTGACAGTGGCCAGGGACCCCTGGGTCCCCAGGCCTGCAGATGTTTCTATGAGGGGAG
AGCTCCTTGGTACATCCATGTGTGGCTCTGCTCCACCCCTGTGCCACCCAGAGCCCTGGGGGGTGGTCTCCATG
CCTGCCACCTGGCATCGGCTTTCTGTGCGCCTCCCACACAAATCAGCCCCAGAAGGCCCGGGGCTTGGCTT
CTGTTTTTTATAAAACACCTCAAGCAGCACTGCAGTCTCCATCTCCTCGTGGGCTAAGCATCACCGCTTCCACG
TGTGTTGTGTTGGTTGGCAGCAAGGCTGATCCAGACCCCTTCTGCCCCACTGCCCTCATCCAGGCCTCTGACCA
GTAGCTGAGAGGGGCTTTTCTAGGCTTCAGAGCAGGGGAGAGCTGGAAGGGGCTAGAAAGCTCCCGCTTGTCT
GTTTCTCAGGCTCCTGTGAGCCTCAGTCTGAGACCAGAGTCAAGAGGAAGTACACGTCCCAATCACCCGTGTCA
GGATTCATCTCAGGAGCTGGGTGGCAGGAGAGGCAATAGCCCCTGTGGCAATTGCAGGACCAGCTGGAGCAGGG
TTGCGGTGTCTCCACGGTGTCTCGCCCTGCCATGGCCACCCAGACTCTGATCTCCAGGAACCCCATAGCCCC
TCTCCACCTACCCCATGTTGATGCCAGGGTCACTCTTGCTACCCGCTGGGCCCCAAACCCCGCTGCCTCTC
TTCTTCCCCCATCCCCACCTGGTTTTGACTAATCCTGCTTCCCTCTCTGGGCTGGCTGCCGGGATCTGGGG
TCCCTAAGTCCCTCTCTTTAAGAACTTCTGCGGGTCAGACTCTGAAGCCGAGTTGCTGTGGGCTGCCCGGAAG
CAGAGCGCCACACTCGTGCTTAAGTCCCCAGCTCTTCCAGAAAACATTAAACTCAGAATTGTGTTTTCAA

372/550

FIGURE 372

MGSRGQGLLLAYCLLLAFASGLVLSRVPHVQGEQQEWEGTEELPSPPDHAERAEQHEKYRPS
QDQGLPASRCLRCCDPGTSMPATAVPQINITILKGEKGDRGDRGLQGYGKTGSAGARGHTG
PKGQKGSMSGAPGERCKSHYAAFSVGRKKPMHSNHYYQTVIFDTEFVNLYDHFNMFTGKFYCYV
PGLYFFSLNVHTWNQKETYLHIMKNEEEVVILFAQVGDRSIMQSQSLMLELREQDQVWVRLYK
GERENAI FSEELDTYITFSGYLVKHATEP

Important features:**Signal sequence.**

amino acids 1-25

N-glycosylation site.

amino acids 93-97

N-myristoylation sites.

amino acids 7-13, 21-27, 67-73, 117-123, 129-135

Amidation site.

amino acids 150-154

Cell attachment sequence.

amino acids 104-107

373/550

FIGURE 373

CGGAGTGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGGCTGCGCTTTCCTGTCCCCAAGCCG
TTCTAGACGCGGGGAAAAATGCTTTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGAA
GCATTTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAATGC
ACCACCATGAGCATCATCACCTACAAGCTCCTAACAAAGAAGATATCTTGAAAATTTAGAGG
ATGAGCGCATGGAGCTCAGTAAGAGCTTTCGAGTATACTGTATTATCCTTGTAACCCAAAG
ATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAACACTGTGACAAAGCAGAGTTCT
TCAGTTCTGAAAATGTTAAAGTGTGAGTCAATTAATATGGACACAAATGACATGTGGTTAA
TGATGAGAAAAGCTTACAAATACGCCTTTGATAAGTATAGAGACCAATACAACCTGGTTCTTCC
TTGCACGCCCCACTACGTTTGCTATCATTGAAAACCTAAAGTATTTTTTGTAAAAAAGGATC
CATCACAGCCTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATATGTGGGTATGG
AAGGAGGAATTGTCTTAAGTGTAGAATCAATGAAAAGACTTAACAGCCTTCTCAATATCCCAG
AAAAGTGTCTGAACAGGGAGGGATGATTTGGAAGATATCTGAAGATAAACAGCTAGCAGTTT
GCCTGAAATATGCTGGAGTATTTGCAGAAAATGCAGAAGATGCTGATGGAAAAGATGTATTTA
ATACCAAATCTGTTGGGCTTTCTATTAAAGAGGCAATGACTTATCACCCCAACCAGGTAGTAG
AAGGCTGTTGTTGAGATATGGCTGTTACTTTTAATGGACTGACTCCAAATCAGATGCATGTGA
TGATGTATGGGGTATACCGCCTTAGGGCATTGCGCATATTTTCAATGATGCATTGGTTTTCT
TACCTCCAAATGGTTCTGACAATGACTGAGAAGTGGTAGAAAAGCGTGAATATGATCTTTGTA
TAGGACGTGTGTTGTCATTATTTGTAGTAGTAACTACATATCCAATACAGCTGTATGTTTCTT
TTTCTTTTCTAATTTGGTGGCACTGGTATAACCACACATTAAAGTCAGTAGTACATTTTTAAA
TGAGGGTGGTTTTTTTCTTTAAACACATGAACATTGTAAATGTGTTGGAAAGAAGTGTTTTA
AGAATAATAATTTTGCAAATAAACTATTAATAAATATTATATGTGATAAATTCTAAATTATGA
ACATTAGAAATCTGTGGGGCACATATTTTGCTGATTGGTTAAAAAATTTTAACAGGTCTTTA
GCGTTCTAAGATATGCAAATGATATCTCTAGTTGTGAATTTGTGATTAAAGTAAACTTTTAG
CTGTGTGTTCCCTTTACTTCTAATACTGATTTATGTTCTAAGCCTCCCCAAGTTCCAATGGAT
TTGCCTTCTCAAAATGTACAATAAGCAACTAAAGAAAATTAAAGTGAAAGTTGAAAAAT

374/550

FIGURE 374

MLSESSSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISEDERMELSKSFRVYCIILV
KPKDVSLWAAVKETWTWKHCDAEFFSSENVKVFESINMDTNDMWLMMRKAYKYAFDKYRDQYNWFFLARPTTFAI
IENLKYFLLKKDPSQPFYLGHTIKSGDLEYVGMEGGIVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAV
CLKYAGVFAENAEDADGKDVFNTKSVGLSIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFG
HIFNDALVFLPPNGSDND

Important features:**Signal sequence:**

amino acids 1-33

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389, 399-403,
409-413, 473-477, 729-733, 748-752**Tyrosine kinase phosphorylation site.**

amino acids 736-743

N-myristoylation site.amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550, 558-564,
651-657, 657-663, 672-672**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

375/550

FIGURE 375

GTTGTGTCCTTCAGCAAAACAGTGGATTTAAATCTCCTTGCACAAGCTTGAGAGCAACACAAT
CTATCAGGAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAAGAAGAAAAAGAAGAGAA
AAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCACGGG
GCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCCCCAA
AGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATTGACAA
CCGGGTACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGACAAGTG
GTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCATCGAGATCCA
GAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAACCACCCAAA
GACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTTGTAGAGATTTCTTCAGATAT
CTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGACCAGAGCCTAC
GGTTACTTGGAGACACATCTCTCCCAAAGCGGTGGCTTTGTGAGTGAAGACGAATACTTGGA
AATTCAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTCCAATGACGTGGC
CGCGCCCGTGGTACGGAGAGTAAAGGTCACCGTGAAGTATCCACCATACATTTCAGAAGCCAA
GGGTACAGGTGTCCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCCTCAGCAGTCCCCTC
AGCAGAATTCCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAAGAAAGGGGTGAAAGT
GGAAACAGACCTTTCCTCTCAAACTCATCTTCTTCAATGTCTCTGAACATGACTATGGGAA
CTACACTTGCGTGCCCTCCAACAAGCTGGGCCACACCAATGCCAGCATCATGCTATTTGGTCC
AGGCGCCGTCAGCGAGGTGAGCAACGGCACGTCGAGGAGGGCAGGCTGCGTCTGGCTGCTGCC
TCTTCTGGTCTTGACCTGCTTCTCAAATTTTGATGTGAGTGCCACTTCCCCACCCGGGAAAG
GCTGCCGCCACCACCACCACCAACACAACAGCAATGGCAACACCGACAGCAACCAATCAGATA
TATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGAAATTTGAGGGAGGGGAACAAA
GAATACTTTGGGGGGAAAGAGTTTTTAAAAAAGAAATTGAAAATTGCCTTGCAGATATTTAGG
TACAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGCACACCCGGCTTGGACCCACTGCA
AGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAAGGGCTCAGCCTCTCTGCCCACAGA
GTGCCCCCACGTGGAACATTCTGGAGCTGGCCATCCCAAATTCATCAGTCCATAGAGACGAA
CAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGGGCACTTTGGTAGACTGTGCCACCACG
CGTGTGTTGTGAAACGTGAAATAAAAAGAGCAAAAAAAA

376/550

FIGURE 376

MKTIQPKMHNSISWAIFTGLAALCLFQGVPPVRSGDATFPKAMDNVTVRQGESATLRCTIDNRV
TRVAWLNIRSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVVDVYDEGPYTCSVQTDNHPKTS
RVHLIVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAVGFVSEDEYLEIQ
GITREQSGDYECASNDVAAPVRRVKVTVNYPPYISEAKGTGVPVGQKGTQCEASAVPSAE
FQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNITCVASNKLGHTNASIMLFGPGA
VSEVSNGTSRRAGCVWLLPLLVLHLLKF

Important features:

Signal peptide:

amino acids 1-28

377/550

FIGURE 377

CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTTGCCCCTTTAGATTGTGAA
ATGTGGCTCAAGGTCTTCACAACCTTCCTTTCCCTTGCAACAGGTGCTTGCTCGGGGCTGAAG
GTGACAGTGCCATCACACACTGTCCATGGCGTCAGAGGTGAGGCCCTCTACCTACCCGTCCAC
TATGGCTTCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCACACA
ATGCCCAAATACTTACTGGGCTCTGTGAATAAGTCTGTGGTTCCCTGACTTGGAATACCAACAC
AAGTTCACCATGATGCCACCCAATGCATCTCTGCTTATCAACCCACTGCAGTTCCTGATGAA
GGCAATTACATCGTGAAGGTCAACATTCAGGGAAATGGAACCTCTATCTGCCAGTCAGAAGATA
CAAGTCACGGTTGATGATCCTGTCAAAAGCCAGTGGTGCAGATTCATCCTCCCTCTGGGGCT
GTGGAGTATGTGGGGAACATGACCCGTGACATGCCATGTGGAAGGGGGCACTCGGCTAGCTTAC
CAATGGCTAAAAAATGGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTTCTCCCCAAAC
AATACCCTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCTGCCTGGTGAGG
AACCCTGTGAGTGAATGGAAAGTGATATCATTATGCCCATCATATATTATGGACCTTATGGA
CTTCAAGTGAATTCTGATAAAGGGCTAAAAGTAGGGGAAGTGTTTACTGTTGACCTGGAGAG
GCCATCCTATTTGATTGTTCTGCTGATTCTCATCCCCCAACACCTACTCCTGGATTAGGAGG
ACTGACAATACTACATATATCATTAAGCATGGGCCTCGCTTAGAAGTGCATCTGAGAAAGTA
GCCCAGAAGACAATGGACTATGTGTGCTTACAAACATAACCGGCAGGCAAGATGAA
ACTCATTTTCAGTTATCATCACTTCGCTAGGACTGGAGAAGCTTGACAGAAAGGAAAATCA
TTGTACCTTTAGCAAGTATAACTGGAATATCACTATTTTTTGATTATATCCATGTGTCTTCTC
TTCTATGGAAAAAATATCAACCCTACAAAGTTATAAAACAGAACTAGAAGGCAGGCCAGAA
ACAGAATACAGGAAAGCTCAAACATTTTCAGGCCATGAAGATGCTCTGGATGACTTCGGAATA
TATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTTCCAGGATTCCAAGCAGGTCTGTTCCA
GCCTCTGATTGTGTATCGGGGCAAGATTTGCACAGTACAGTGTATGAAGTTATTCAGCACATC
CCTGCCCAGCAGCAAGACCATCCAGAG**TGA**ACTTTCATGGGCTAAACAGTACATTCGAGTGAA
ATTCTGAAGAAACATTTTAAGGAAAAACAGTGGAAAAGTATATTAATCTGGAATCAGTGAAGA
AACCAGGACCAACACCTCTTACTCATTATTCTTTACATGCAGAATAGAGGCATTTATGCAAA
TTGAACTGCAGGTTTTTTCAGCATATACACAATGTCTTGTGCAACAGAAAAACATGTTGGGGAA
ATATTCCTCAGTGGAGAGTCGTTCTCATGCTGACGGGGAGAACGAAAGTGACAGGGGTTTCCT
CATAAGTTTTGTATGAAATATCTCTACAAACCTCAATTAGTTCTACTCTACACTTTCACTATC
ATCAACACTGAGACTATCCTGTCTCACCTACAAATGTGGAACTTTACATTGTTTCGATTTTTTC
AGCAGACTTTGTTTTATTAAATTTTTATTAGTGTTAAGAATGCTAAATTTATGTTTCAATTTT
ATTTCCAAATTTCTATCTTGTTATTTGTACAACAAAGTAATAAGGATGGTTGTCACAAAAACA
AACTATGCCTTCTCTTTTTTTTCAATCACCAGTAGTATTTTTTGAGAAGACTTGTGAACACTT
AAGGAAATGACTATTAAAGTCTTATTTTTATTTTTTCAAGGAAAGATGGATTCAAATAAATT
ATTCTGTTTTTGCTTTTAAAAAAAAAAAAAA

378/550

FIGURE 378

MWLKVFTTFLSFATGACSGLKVTVPSTVHGVGRQALYLPVHYGFHTPASDIQIIWLFERPHTMPKYLLGSVNKS
VVPDLEYQHKFTMMPPNASLLINPLQFPDEGNYIVKVNIQNGTSLASQKIQVTVDVDPVTKPVVQIHPPSGAVEY
VGNMTLTCHVEGGTRLAYQWLKNGRPVHTSSTYSFSPQNNTLHIAPVTKEDIGNYSCLVRNPVSEMESDIIMPII
YYGPYGLQVNSDKGLKVGVEFTVDLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTMD
YVCCAYNNITGRQDETHFTVIITSVGLEKLAQKGKSLSPASITGISLFLIISMCLLFLWKKYQPYKVIKQKLEG
RPETEYRKAQTFSGHEDALDDFGIYEFVAFPDVSGVSRIPSRVPASDCVSGQDLHSTVYEVIQHIPAQQQDHPE

Important features:**Signal sequence:**

amino acids 1-18

Transmembrane domain:

amino acids 341-359

N-glycosylation site.

amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208, 276-280, 308-312

Casein kinase II phosphorylation site.

amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

Tyrosine kinase phosphorylation site.

amino acids 272-280

N-myristoylation site.

amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237, 239-245

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

379/550

FIGURE 379

ATAGTAGAAGAATGTCTCTGAAATTACTGGATGAGTTTCAGTCATACTTTCACATGGGCACAA
TTTCACATTCAAGCTCCTTATCCTAGGCTAATTTTATATTATGTTAAATCACTTGTTTTTGT
CTCACGGCTTCCTGCCTGCTATAGGCATAATTACGAGGAAGCAGAACTTCTCCAGAAGCAAGC
GCACATGCGTTCCAAAATAAGAGCAAATTCGCTCTAAACACAGGAAAAGACCTGAAGCTTTAA
TTAAGGGGTTACATCCAACCCAGAGCGCTTTTGTGGGCACTGATTGCTCCAGCTTCTGCGTC
ACTGCGCGAGGGAAGAGGGAAGAGGATCCAGGCGTTAGACATGTATAGACACAAAAACAGCTG
GAGATTGGGCTTAAAATACCCACCAAGCTCCAAAGAAGAGACCCAAGTCCCCAAAACATTGAT
TTCAGGGCTGCCAGGAAGGAAGAGCAGCAGCAGGGTGGGAGAGAAGCTCCAGTCAGCCCACAA
GATGCCATTGTCCCCGGCCTCCTGCTGCTGCTGCTCTCCGGGGCCACGGCCACCGCTGCCCT
GCCCCTGGAGGGTGGCCCCACCGGCCGAGACAGCGAGCATATGCAGGAAGCGGCAGGAATAAG
GAAAAGCAGCCTCCTGACTTTCCTCGCTTGGTGGTTTGAGTGGACCTCCCAGGCCAGTGCCGG
GCCCCTCATAGGAGAGGAAGCTCGGGAGGTGGCCAGGCGGCAGGAAGGCGCACCCCCCAGCA
ATCCGCGCGCCGGGACAGAATGCCCTGCAGGAACCTTCTTCTGGAAGACCTTCTCCTCCTGCAA
ATAG

380/550

FIGURE 380

MYRHKNSWRLGLKYPPSSKEETQVPKTLISGLPGRKSSSRVGEKLQSAHKMPLSPGLLLLLLS
GATATAALPLEGGPTGRDSEHMQEAAGIRKSSLLTFLAWWFEWTSQASAGPLIGEEAREVARR
QEGAPPQQSARRDRMPCRNFFWKTFSSCK

Important features:

Transmembrane domain:

amino acids 51-69

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 35-39, 92-96

N-myristoylation sites.

amino acids 64-70, 75-81, 90-96

Amidation site.

amino acids 33-37

381/550

FIGURE 381

GGCGCCGGTGCACCGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCCTGCGCGCCCCGGCCCGCC
GCGCCGCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCCGGGCCCGCGCCCCGCGC
CCGCGCCCAAGGTGAGCGCTCCGCCCCGCGCGAGGCCCGCCCCGGCCCGCCCCCGCCCCGCC
CGGCCGGCGGGGAACCGGGCGGATTCTCGCGCGTCAAACCACCTGATCCCATAAACATTC
ATCCTCCCCGGCGGCCCGCGCTGCGAGCGCCCCGAGTCCGCGCCCGCCCGCCCTCGCCCTG
TGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCAGAGCCGGGCGGAGCGGAGCG
CGCCGAGCCTCGTCCCGCGGCCGGGCCGGGCCGTAGCGGCGGCGCCTGGATGCGGAC
CCGGCCGCGGGGAGACGGGCGCCCCGCCCGAAACGACTTTCAGTCCCCGACGCGCCCCGCCA
ACCCCTACGATGAAGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTGCTGTGGCTG
CAGGCCTGGCAGGTGGCAGCCCCATGCCAGGTGCCTGCGTATGCTACAATGAGCCCAAGTG
ACGACAAGCTGCCCCAGCAGGGCCTGCAGGCTGTGCCGTGGGCATCCCTGCTGCCAGCCAG
CGCATCTTCTGCACGGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTCCGTGCTGCCG
AACCTACCATCCTGTGGCTGCACTCGAATGTGCTGGCCCGAATTGATGCGGCTGCCTTCACT
GGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCCGGTCTGTGGACCCT
GCCACATTCCACGGCCTGGGCCGCCTACACACGCTGCACCTGGACCGCTGCGGCCTGCAGGAG
CTGGGCCCGGGGCTGTTCCGCGGCCTGGCTGCCCTGCAGTACCTCTACCTGCAGGACAACGCG
CTGCAGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCACACACCTCTTCTGCAC
GGCAACCGCATCTCCAGCGTGCCCGAGCGCGCCTTCCGTGGGCTGCACAGCCTCGACCGTCTC
CTACTGCACCAGAACCGCGTGCCCGATGTGCACCCGCATGCCTTCCGTGACCTTGGCCGCCCTC
ATGACACTCTATCTGTTTGCCAACAATCTATCAGCGCTGCCCACTGAGGCCCTGGCCCCCCTG
CGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTGTGTGACTGCCGGGCACGCCCA
CTCTGGGCCTGGCTGCAGAAGTTCCGCGGCTCCTCCTCCGAGGTGCCCTGCAGCCTCCCGCAA
CGCCTGGCTGGCCGTGACCTCAAACGCCTAGCTGCCAATGACCTGCAGGGCTGCGCTGTGGCC
ACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCACCGATGAGGAGCCGCTGGGGCTTCCC
AAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCTCAGTACTGGAGCCTGGAAGACCAGCTTCG
GCAGGCAATGCGCTGAAGGGACGCGTGCCGCCCGGTGACAGCCCGCCGGGCAACGGCTCTGGC
CCACGGCACATCAATGACTCACCCTTTGGGACTCTGCCTGGCTCTGCTGAGCCCCGCTCACT
GCAGTGCGGCCCCGAGGGCTCCGAGCCACCAGGGTTCCCCACCTCGGGCCCTCGCCGGAGGCCA
GGCTGTTACGCAAGAACCGCACCCGCAGCCACTGCCGTCTGGGCCAGGCAGGCAGCGGGGGT
GGCGGGACTGGTGACTCAGAAGGCTCAGGTGCCCTACCCAGCCTCACCTGCAGCCTCACCCCC
CTGGGCCTGGCGCTGGTGCTGTGGACAGTGCTTGGGCCCTGCTGAACCCCCAGCGGACACAAGA
GCGTGCTCAGCAGCCAGGTGTGTGTACATACGGGGTCTCTCTCCACGCCGCCAAGCCAGCCGG
GCGGCCGACCCGTGGGGCAGGCCAGGTCCTCCCTGATGGACGCCTGCCGCCCGCCACC
CCCATCTCCACCCCATCATGTTTACAGGGTTTCGGCGGCAGCGTTTGTTCAGAACGCCGCCTC
CCACCCAGATCGCGGTATATAGAGATATGCATTTTATTTTACTTGTGTAAAAATATCGGACGA
CGTGGAATAAAGAGCTCTTTTCTTAAAAAA

382/550

FIGURE 382

MKRASAGGSRL LAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRIF
LHG NRISHVPAASFRACRNLTILWLHSNVLARIDAAAF TGLALLEQLDLSDNAQLRSVDPATF
HGLGRLHTLHLDRCLQELGPGLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLHG NR
ISSVPERAFRGLHSLDRLLLHQNRVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAPLRAL
QYLRLNDNPWVCD CRARPLWAWLQKFRGSSSEVP CSLPQRLAGRDLKRLAANDLQGCAVATGP
YHPIWTGRATDEEPLGLPKCCQPDAA DKASVLEPGRPASAGNALKGRVPPGDSPPGN SGPRH
INDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQAGSGGGGT
GDSESGALPSLTCSLTPLGLALVLWTVLGPC

Important features:**Signal peptide:**

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

383/550

FIGURE 383

TTCGTGACCCCTTGAGAAAAGAGTTGGTGGTAAATGTGCCACGTCTTCTAAGAAGGGGGAGTCCTGAACTTGTCTG
AAGCCCTTGTCCGTAAGCCTTGAAC TACGTTCTTAAATCTATGAAGTCGAGGGACCTTTTCGCTGCTTTTGTAGGG
ACTTCTTTTCCTTGCTTCAGCAAC ATGAGGCTTTTCTTGTGGAACGCGGTCTTGACTCTGTTTCGTCACCTTCTTTGA
TTGGGGCTTTGATCCCTGAACCAGAAGTGAAAATTGAAGTTCTCCAGAAGCCATTTCATCTGCCATCGCAAGACCA
AAGGAGGGGATTTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTTCACTCCACTCACA
AACATAACAATGGTCAGCCCATTGGTTTACCCTGGGCATCCTGGAGGCTCTCAAAGGTTGGGACCAGGGCTTGA
AAGGAATGTGTGTAGGAGAGAAGAGAAAGCTCATCATTCCTCCTGCTCTGGGCTATGGAAAAGAAGGAAAAGGTA
AAATCCCCCAGAAAGTACACTGATATTTAATATTGATCTCCTGGAGATTGCGAAATGGACCAAGATCCCATGAAT
CATTCCAAGAAATGGATCTTAATGATGACTGGAACTCTCTAAAGATGAGGTTAAAGCATATTTAAAGAAGGAGT
TTGAAAACATGGTGCGGTGGTGAATGAAAGTCATCATGATGCTTTGGTGGAGGATATTTTTGATAAAGAAGATG
AAGACAAAGATGGGTTTATATCTGCCAGAGAATTTACATATAAACACGATGAGTTA TAGAGATACATCTACCCTT
TTAATATAGCACTCATCTTTCAAGAGAGGGCAGTCATCTTTAAAGAACATTTTATTTTTTATACAATGTTCTTTCT
TGCTTTGTTTTTTATTTTATATATTTTTTCTGACTCCTATTTAAAGAACCCCTTAGGTTTCTAAGTACCCATTT
CTTTCTGATAAGTTATTGGGAAGAAAAAGCTAATTGGTCTTTGAATAGAAGACTTCTGGACAATTTTTCACTTTC
ACAGATATGAAGCTTGTGTTTACTTTCTCACTTATAAATTTAAATGTTGCAACTGGGAATATACCACGACATGA
GACCAGGTTATAGCACAAATTAGCACCCCTATATTTCTGCTTCCCTCTATTTTCTCCAAGTTAGAGGTCAACATTT
GAAAAGCCTTTTGAATAGCCCAAGGCTTGCTATTTTCATGTTATAATGAAATAGTTTATGTGTAAGTGGCTCTG
AGTCTCTGCTTGAGGACCAGAGGAAAATGGTTGTTGGACCTGACTTGTTAATGGCTACTGCTTTACTAAGGAGAT
GTGCAATGCTGAAGTTAGAAACAAGGTTAATAGCCAGGCATGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGA
GGCTGAGGCGGGCGGATCACCTGAGGTTGGGAGTTCGAGACCAGCCTGACCAACACGGAGAAACCTATCTCTAC
TAAAAATACAAAGTAGCCCGCGTGGTGATGCGTGCTGTAATCCCAGCTACCCAGGAAGGCTGAGGCGGCAGAA
TCACTTGAACCCGAGGCCGAGGTTGCGGTAAGCCGAGATCACCTNCAGCCTGGACACTCTGTCTCGAAAAAAGAA
AAGAACACGGTTAATACCATATNAATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGG
CTCCTAGTGATTGGTGGCCTATTATGATAAATAGGACAAATCATTTATGTGTGAGTTTCTTTGTAATAAAATGTA
TCAATATGTTATAGATGAGGTAGAAAGTTATATTTATATTCAATATTTACTTCTTAAGGCTAGCGGAATATCCTT
CCTGGTTCCTTAATGGGTAGTCTATAGTATATTATACTACAATAACATTGTATCATAAGATAAAGTAGTAAACCA
GTCTACATTTTCCCATTTCTGTCTCATCAAAAAGTGAAGTTAGCTGGGTGTGGTGGCTCATGCCTGTAATCCCAG
CACTTTGGGGGCCAAGGAGGTGGATCACTTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCT
TGTCTCTACTAAAAATACAAAAATTAGCCAGGCGTGGTGGTGCACACCTGTAGTCCCAGCTACTCGGGAGGCTGA
GACAGGAGATTTGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGTGCCACTGCACTCCAGCCTGGG
TGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAAGAAGCAGACCTACAGCAGCTACTATTGAATAAATACCTA
TCCTGGATTTT

384/550

FIGURE 384

MRLFLWNAVLTSLFVTSLLIGALIPPEVKIEVLQKPFICHRKTKGGDLMLVHYEGYLEKDGSLE
HSTHKHNNNGQPIWFTLGILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGKIPPESTLI
FNIDLLEIRNGPRSHESFQEMDLNDDWKLKSKDEVKAYLKKEFEKHGAVVNESHHDALVEDIFD
KEDEKDKDGFISAREFTYKHDEL

Important features:**Signal peptide:**

amino acids 1-20

N-glycosylation site.

amino acids 176-179

Casein kinase II phosphorylation site.

amino acids 143-146, 156-159, 178-181 and 200-203

Endoplasmic reticulum targeting sequence.

amino acids 208-211

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 78-114 and 118-131

EF-hand calcium-binding domain.

amino acids 191-203, 184-203 and 140-159

S-100/ICaBP type calcium binding domain

amino acids 183-203

385/550

FIGURE 385

CTCCACGGTGTCCAGCGCCAGAAATGCGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCTC
CCAGGTTATGAAGCCCTGGAGGGGCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGTCC
CTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGTGGG
ATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAATGAAG
GGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGAACCTC
ACCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAGTCTTTA
CTGATCTCTCTGTTCGTCTTTCCAGGACCCTGCTGTCTCCTCCCTCCCCTTCTCCACCTTCCAG
CCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAACCCAGCCCCCAGGATTG
ACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGGGCTGAGGCC
CCTCCATTGCCAGGGACTTCCAGTACGGGCACGAAAGGACTTCTCAGTACACAGGAACCTCT
CCTCACCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCATGCAGCTGGACTCCACC
TCAGCAGAGGACACCACTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGGGTGTCCATCCCG
ATGGTCCGCATACTGGCCCCAGTCCCTGGTGTCTGTGAGCCTTCTGTTCAGCCGCAGGCCTGATC
GCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAAGCTCAACAGGCCACGGAGACACAGAGG
AACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCCCCCTTCCAGGCCCTT
GAGGGGGACGTGATCTCGATGCCTCCCTCCACACATCTGAGGAGGAGCTGGGCTTCTCGAAG
TTTGTCTCAGCGTAGGGCAGGAGGCCCTCCTGGCCAGGCCAGCAGTGAAGCAGTATGGCTGGC
TGGATCAGCACCGATTCCCGAAAGCTTTCCACCTCAGCCTCAGAGTCCAGCTGCCCGGACTCC
AGGGCTCTCCCCACCCTCCCCAGGCTCTCCTCTTGATGTTCCAGCCTGACCTAGAAGCGTTT
GTCAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTGGAGACTGGGACATCCCTGAT
AGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCAGCAGGGCCAGACAAGGCTCAG
TGGATCTGGTCTGAGTTTCAATCTGCCAGGAACCTCCTGGGCCTCATGCCAGTGTCGGACCCT
GCCTTCTCTCCACTCCAGACCCACCTTGTCTTCCCTCCCTGGCGTCTCAGACTTAGTCCCA
CGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCTGGGGTGAGACTGGGATTCTGGCTT
CTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCTGTGAAAAACGTGATTCTGGCCCCA
CCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAGGACTCTGAATTCTAACAATGCCAGT
GACTGTGCACTTGAGTTTGAAGGGCAGTGGGCCTGATGAACGCTCACACCCCTTCAGCTTAG
AGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCCCAATAGATCTGCTCTGTCTGCGACACCA
GATCCACGTGGGGACTCCCCTGAGGCCTGCTAAGTCCAGGCCTTGGTCAGGTCAGGTGCACAT
TGCAGGATAAGCCCAGGACCGGCACAGAAGTGGTTGCCTTTNCCATTTGCCCTCCCTGGNCCA
TGCCTTCTTGCCCTTTGGAAAAAATGATGAAGAAAACCTTGGCTCCTTCTTGTCTGGAAAGGG
TTACTTGCCTATGGGTTCTGGTGGCTAGAGAGAAAAGTAGAAAACAGAGTGCACGTAGGTGT
CTAACACAGAGGAGTAGGAACAGGGCGGATACCTGAAGGTGACTCCGAGTCCAGCCCCCTG
GAGAAGGGGTGCGGGGTGGTGGTAAAGTAGCACTACTATTTTTTTTCTTTTCCATTATT
ATTGTTTTTTAAGACAGAATCTCGTGCTGCTGCCAGGCTGGAGTGCAGTGGCACGATCTGCA
AACTCCGCCTCCTGGGTTCAAGTGATTCTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAG
GCACGCACCACACCTGGCTAATTTTTGTACTTTTAGTAGAGATGGGGTTTACCATGTTG
GCCAGGCTGGTCTTGAACCTCTGACCTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCG
GGATTACAGGCATGAGCCACTGTGTCTGGCCCTATTTCTTTAAAAAGTGAATTAAGAGTTG
TTCAGTATGCAAACTTGGAAGATGGAGGAGAAAAAGAAAAGGAAGAAAAAATGTCACCCA
TAGTCTACCAAGACTATCATTATTTCTGTTTGTGTAATCTTGTATATACAACCTTTGTATCCTGC
ACATAATTTGCCGGTGTCTTTTACAGAGCAATTATCTTGTATATACAACCTTTGTATCCTGC
CTTTTCCACCTTATCGTTCCATCACTTTATTCAGCACTTCTCTGTGTTTTACAGACCTTTTT
ATAAATAAAATGTTTCATCAGCTGCATAAAAAAAAAAAAAA

386/550

FIGURE 386

MRLLVLLWGCLLLPGYEALLEGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCSG
TIYAE EEGQETMKGRVSIRD SRQELSLIVTLWNLTLQDAGEYWCGVEKRGPD E SLLISLFVFP
GPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPGTSQ
YGHERTSQYTGTSPHPATSPPPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRILAPV
LVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAE E KEAPSQAPEGDVISMP
PLHTSEEELGFSKFVSA

Important features:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

387/550

FIGURE 387

GCGCCGGGAGCCCATCTGCCCCAGGGGCACGGGGCGCGGGGCCGGCTCCCGCCCGGCACATG
GCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCGCGCCAGCTCGCCCGAGGTCGGTCCGGAGG
CGCCCGGCCGCCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCTCCGGGGATCGGG
ATGTCCCTCCTCCTTCTCCTCTTGCTAGTTTCTACTATGTTGGAACCTTGGGGACTCACACT
GAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGCTTCCA
GAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAAGTGGTG
ATCACTTACTCCAGTCGTCATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCGAGTGGCC
TTTGCTTCCAATTTCTGGCAGGAGATGCCTCCTTGACAGATTGAACCTCTGAAGCCCAGTGAT
GAGGGCCGGTACACCTGTAAGGTTAAGAATTGAGGGCGCTACGTGTGGAGCCATGTCATCTTA
AAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGACAGAAGGAAGT
GACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTTGTGTATTACTGGCAGCGA
ATCCGAGAGAAAAGGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATTGACTACAACCAC
CCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTACAGTGCACAGCA
GGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGGAGTAAGTGTACAGTATGTACAAAGCATC
GGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTGATTTTCTCTTGGTG
TGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGAGAGACCTAATGAAATT
CGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCTCCTCTTCCCTCAGGCTCT
CGGAGCTCACGCTCTGGTTCTTCCCTCCACTCGCTCCACAGCAAATAGTGCCTCACGCAGCCAG
CGGACACTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCACCCAGGCATACAGCCTAGTG
GGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATGCTAATCTGACCAAAGCAGAA
ACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAAACGGTCT**TGA**ATTACAATGGAC
TTGACTCCACGCTTTTCTAGGAGTCAGGGTCTTTGGACTCTTCTCGTCATTGGAGCTCAAGT
CACCAGCCACACAACCAGATGAGAGGTGATCTAAGTAGCAGTGAGCATTGCACGGAACAGATT
CAGATGAGCATTTTCTTATACAATACCAAACAAGCAAAAGGATGTAAGCTGATTCATCTGTA
AAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGGAAAGCAGGAGTCCAAATCTATTTGT
TGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTGAGGTGAATATACCTAAACCTTTTAAT
GTGGGATATTTTGTATCAGTGCTTTGATTACAAATTTTCAAGAGGAAATGGGATGCTGTTTGT
AAATTTTCTATGCATTTCTGCAAACTTATTGGATTATTAGTTATTCAGACAGTCAAGCAGAAC
CCACAGCCTTATTACACCTGTCTACACCATGTACTGAGCTAACCCTTCTAAGAACTCCAAA
AAAGGAAACATGTGTCTTCTATTCTGACTTAACTTCATTTGTCATAAGGTTTGGATATTAATT
TCAAGGGGAGTTGAAATAGTGGGAGATGGAGAAGAGTGAATGAGTTTCTCCCACTCTATACTA
ATCTCACTATTTGTATTGAGCCCAAATAACTATGAAAGGAGACAAAATTTGTGACAAAGGA
TTGTGAAGAGCTTCCATCTTCATGATGTTATGAGGATTGTTGACAAACATTAGAAATATATA
ATGGAGCAATTGTGGATTTCCCTCAAATCAGATGCCTCTAAGGACTTTCCTGCTAGATATTT
CTGGAAGGAGAAAATACAACATGTCATTTATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAA
AAGGGATCTAGGAATGCTGAAAGATTACCCAACATACCATTATAGTCTCTTCTTTCTGAGAAA
ATGTGAAACCAGAATTGCAAGACTGGGTGGACTAGAAAGGGAGATTAGATCAGTTTCTCTTA
ATATGTCAAGGAAGGTAGCCGGGCATGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAG
GTTGCAGTGAGCCGAGATTATGCCATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

388/550

FIGURE 388

MSLLLLLLLVSYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV
ITYSSRHVYNNLTTEEQGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHVIL
KVLVRPSKPKCELEGELETEGSDTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRIDYNH
PGRVLLQNLMTSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLIFLLV
WLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANSASRSQ
RTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Important features:**Signal sequence:**

amino acids 1-16

Transmembrane domain:

amino acids 232-251

389/550

FIGURE 389

GCGGCACCTGGAAG**ATG**CGCCCATTTGGCTGGTGGCCTGCTCAAGGTGGTGTTCGTGGTCTTCG
CCTCCTTGTGTGCCTGGTATTCGGGGTACCTGCTCGCAGAGCTCATTCCAGATGCACCCCTGT
CCAGTGCTGCCTATAGCATCCGCAGCATCGGGGAGAGGCCTGTCCTCAAAGCTCCAGTCCCCA
AAAGGCAAAAATGTGACCACTGGACTCCCTGCCCATCTGACACCTATGCCTACAGGTTACTCA
GCGGAGGTGGCAGAAGCAAGTACGCCAAAATCTGCTTTGAGGATAACCTACTTATGGGAGAAC
AGCTGGGAAATGTTGCCAGAGGAATAAACATTGCCATTGTCAACTATGTAAGTGGGAATGTGA
CAGCAACACGATGTTTTGATATGTATGAAGGCGATAACTCTGGACCGATGACAAAGTTTATTC
AGAGTGCTGCTCCAAAATCCCTGCTCTTCATGGTGACCTATGACGACGGAAGCACAAGACTGA
ATAACGATGCCAAGAATGCCATAGAAGCACTTGGAAGTAAAGAAATCAGGAACATGAAATTCA
GGTCTAGCTGGGTATTTATTGCAGCAAAGGCTTGGAAGTCCCTTCCGAAATTCAGAGAGAAA
AGATCAACCACTCTGATGCTAAGAACAACAGATATTCTGGCTGGCCTGCAGAGATCCAGATAG
AAGGCTGCATACCCAAAGAACGAAGCT**TGA**CACTGCAGGGTCCTGAGTAAATGTGTTCTGTATA
AACAAATGCAGCTGGAATCGCTCAAGAATCTTATTTTTCTAAATCCAACAGCCCATATTTGAT
GAGTATTTTGGGTTTGTGTAAACCAATGAACATTTGCTAGTTGTATCAAATCTTGGTACGCA
GTATTTTTTATACCAGTATTTTATGTAGTGAAGATGTCAATTAGCAGGAACTAAAATGAATGG
AAATTCTTAAAAAAAAAA

390/550

FIGURE 390

MRPLAGGLLKVVVFVVFASLCAWYSGYLLAELIPDAPLSSAAYSIRSIGERPVLKAPVPKRQKC
DHWTPCPSDTYAYRLLSGGGRSKYAKICFEDNLLMGEQLGNVARGINIAIVNYVTGNVTATRC
FDMYEGDNSGPMTKFIQSAAPKSLLFMVTYDDGSTRLNNDKNAIEALGSKEIRNMKFRSSWV
FIAAKGLELPSEIQREKINHSDAKNNRYSGWPAEIQIEGCIPKERS

Important features:

Signal sequence.

amino acids 1-20

N-glycosylation sites.

amino acids 120-124, 208-212

Glycosaminoglycan attachment site.

amino acids 80-84

N-myristoylation sites.

amino acids 81-87, 108-114, 119-125

391/550

FIGURE 391

GGGGGCTTTCTTGGGCTTGGCTGCTTGGAAACACCTGCCTCCAAGGACCGGCCTCGGAGGGGTGCGCGGGAAAGGG
AGGGAAGAAGGAAGGGCGGGGCCCCCTGCGCCCCCGCGCCTCTGCGCGCCCCGTGTCGCCCCGGCCCC
AGCCCCAGCCAGCCCCGCGGGCCGGTCACACGCGCAGCCAGCCGCGCCCTCCCGCGCCCAAGCGCGCCGCTCTG
CTGTGCCCTGCGCCCTTGCCCCGCGCCAGCTTCTGCGCCCGCAGCCCGCCCGCGCCCCGGTGACCGTGACCCCT
GCCCTGGGCGCGGGGCGGAGCAGGCATGTCGCCCCGGGACCGCTACCCAGCGCTGGCCCTGGTGCTCCTGGC
AGTGACCCTGGCCGGGGTTCGGAGCCAGGGCGCAGCCCTCGAGGACCCTGATTATTACGGGCAGGAGATCTGGAG
CCGGGAGCCCTACTACGCGCGCCCGGAGCCCGAGCTCGAGACCTTCTCTCCGCGCTGCCTGCGGGGCCCCGGGA
GGAGTGGGAGCGGCGCCCGCAGGAGCCAGGCCGCCCAAGAGGGGCCACCAAGCCCAAGAAAGCTCCCAAGAGGGA
GAAGTCGGCTCCGGAGCCGCTCCACCAGGTAAACACAGCAACAAAAAGTTATGAGAACCAAGAGCTCTGAGAA
GGCTGCCAACGATGATCACAGTGTCCGTGTGGCCCGTGAAGATGTCAGAGAGAGTTGCCACCTCTTGGTCTGGA
AACCTTAAAAATCACAGACTTCCAGCTCCATGCCTCCACGGTGAAGCGCTATGGCCTGGGGGCACATCGAGGGAG
ACTCAACATCCAGGCGGGCATTAAATGAAAATGATTTTTATGACGGAGCGTGGTGCGCGGGAAGAAATGACCTCCA
GCAGTGGATTGAAGTGGATGCTCGGCGCTGACCAAGATTCACTGGTGTCATCACTCAAGGGAGGAATCCCTCTG
GCTGAGTGACTGGGTGACATCCTATAAGGTCATGGTGAGCAATGACAGCCACACGTGGGTCACTGTTAAGAATGG
ATCTGGAGACATGATATTGAGGGAACAGTGAGAAGGAGATCCCTGTTCTCAATGAGCTACCCGTCCCCATGGT
GGCCCGCTACATCCGCATAAACCCCTCAGTCTGGTTTGATAATGGGAGCATCTGCATGAGAAATGGAGATCCTGGG
CTGCCCACTGCCAGATCCTAATAATTATTATCACCGCCGGAACGAGATGACCACCACTGATGACCTGGATT'TTAA
GCACCACAATTATAAGGAAATGCGCCAGTTGATGAAAGTTGTGAATGAAATGTGTCCCAATATACCAGAATTTA
CAACATTGGAAGAAAGCCAGGCGCTGAAGCTGTATGCTGTGGAGATCTCAGATCACCCCTGGGGAGCATGAAGT
CGGTGAGCCCGAGTTCACCTACATCGCGGGGGCCACGGCAATGAGGTGCTGGGCGGGAGCTGCTGCTGCTGCT
GGTGACGTTTCGTGTGTGAGGAGTACTTGGCCCGGAATGCGCGCATCGTCCACCTGGTGGAGGAGACGCGGATTCA
CGTCTCTCCCTCCCTCAACCCCGATGGCTACGAGAAGGCCCTACGAAGGGGGCTCGGAGCTGGGAGGCTGGTCCCT
GGGACGCTGGACCCACGATGGAATTGACATCAACAACACTTCTGATT'TAAACACGCTGCTCTGGGAGGCAGA
GGATCGACAGAATGTCCCCAGGAAAGTTCCCAATCACTATATTGCAATCCCTGAGTGGTTTCTGTGCGAAAATGC
CACGGTGGCTGCCGAGACCAGAGCAGTCATAGCCTGGATGGAAGAAATCCCTTTTGTGCTGGGCGGCAACCTGCA
GGCGCGCGAGCTGGTGGTGGCGTATCCCTACGACCTGGTGGCGTCCCCCTGGAAGACGCGAGGAACACACCCCCAC
CCCCGATGACCAGTGTTCGCTGGCTGGCCTACTCCTATGCCTCCACACACCGCCTCATGACAGACGCCCCGGAG
GAGGGTGTGCCACCGGAGGACTTCCAGAAGGAGGAGGCCACTGTCAATGGGGCCCTCTGGCACACCGCTCGCTGG
AAGTCTGAACGATTTCACTACCTTCATACAACTGCTTCAACTGTCCATCTACGTGGGCTGTGATAAATACCC
ACATGAGAGCCAGCTGCCCGAGGAGTGGGAGAATAACCGGAATCTCTGATCGTGTTCATGGAGCAGGTTTCATCG
TGGCATTAAAGGCTTGGTGAGAGATTCACATGGAAGGAATCCCAACGCCATTATCTCCGTAGAAGGCATTAA
CCATGACATCCGAACAGCCAACGATGGGGATTACTGGCGCCTCCTGAACCTGGAGAGTATGTGGTCACAGCAAA
GGCCGAAGGTTTCACTGCATCCACCAAGAAGTGTATGGTTGGCTATGACATGGGGGCCACAAGGTGTGACTTCAC
ACTTAGCAAAACCAACATGGCCAGGATCCGAGAGATCATGGAGAAGTTTGGGAAGCAGCCCGTCAGCCTGCCAGC
CAGGCGGCTGAAGCTGCGGGGGCGAAGAGACGACAGCGTGGGTGAACCTCCTGGGCCCTGAGACTCGTCTGGG
ACCCATGCAAAATTAACCAACCTGGTAGTAGCTCCATAGTGGACTCACTCACTGTTGTTTCTCTGTAATTCAAG
AAGTGCCTGGAAGAGAGGGTGCATTGTGAGGCAGGTCCCAAGGGAAGGCTGGAGGCTGAGGCTGTTTTCTTTT
CTTTGTTCCCATTTATCCAAATAACTTGGACAGAGCAGCAGAGAAAAGCTGATGGGAGTGAGAGAATCAGCAAG
CCAACCTGGGAATCAGAGAGAGAAGGAGAAGGAGGGGAGCCTGTCCGTTGAGAGCCTCTGGCTGCATAGAAAAGG
ATTCTGGTGTCTCCCTGTTTGGCTGGCAGCAAGGGTTCACGTGCATTTGCAATTTGCACAGCTAAAATTGCAG
CATTTCCCGAGCTGGGCTGTCCCAATGTTACCATTTGAGATGCTCCAGGCGTCTAAGAGAATCCACCTCTC
TGGCCCTGGGACATTGCAAGCTGCTACAAATAAATTTCTGTGTTCTTTTGACAATAGCGTCATTGCCAAGTGCACA
TCAGTGAGCCTCTTGAATCTGTTTAGTCTCCTTTTCAACAAAGGAGTGTGTTGAGAAAAGGAGAGAGAGGCTGA
GATCATTGAGGAGTTTGTGGGCGAGCAAGCATGGAGCTTCTTGACAAAATTTCTGGGTCCATAAACACCCCCAAA
GTCCCTGCTGATCCAGTAGCCCTGGAGGTTCCCGAGGTAGGAGAGCCAGAGGTGCCAGCCTTCTGAAAGGGCCA
GAAAATTTAGCCTGGATCTCCTCTTTTACCTGCTAGGACTGGAAAGAGCCAGAAGTGGGGTGGCCTGAAGCCCTC
TCTCTGCTTGAAGTATTGCCCTGTGTGGAATTGAGTGCTCATGGGTGGCCTCATATCAGCCTGGGAGTTATTT
TTGATATGTAGAATGCCAGATCTCCAGATTAGGCTAAATGTAATGAAAACCTCTAGGATTATCTGTGGAGCAT
CAGTTTGGGAAGAATTATTGAATTATCTTGCAAGAAAAAGTATGTCTCACTTTTTGTAAATGTTGCTGCTCAT
TGACCTGGGAAAAATGAAAAAATAAAGCAATGGTAAGACCTTAAAAAATAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

392/550

FIGURE 392

MSRPGTATPALALVLLAVTLAGVGAQGALEDPDYYGQEIWSREPYYARPEPELETFSPPLPA
GPGEWEERRPQEPRPPKRATKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHSVR
VAREDVRESCPPLGLETLKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCAGR
DLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSTWVTVKNGSGDMIFEGNS
EKEIPVLNELPVPVMVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTTTDDLD
FKHHNYKEMRQLMKVVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPGEHEVGEPEFHYIAGA
HGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLP SLNPDGYEKAYEGGSELGGWS
LGRWTHDGIDINNNFPDLNTLLWEAEDRQNVPRKVPNHYYIAIPEWFLSENATVAAETRAVIAW
MEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYSASTHRLMTDARR
RVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHESQLPEEWENNR
ESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRLNPG EYVVTAKA
EGFTASTKNCMVGYDMGATRCDFTL SKTNMARI REIMEKFGKQPVSLPARRLKLGRKRRQRG

393/550

FIGURE 393

GTCCACATCCTGCTCAACTGGGTCAGGTCCCTCTTAGACCAGCTCTTGTCATCATTTGCTGAAGTGGACCAAC
TAGTTCCCAAGTAGGGGCTCTCCCTGGCAATTCTTGATCGGCGTTTGGACATCTCAGATCGCTTCCAATGAAGA
TGGCCTTGCTTGGGCTGCTGCTTGTTCATAATCATCTAATATGGGACAAGGTTGTGCCGGCAGCTCTGGGG
AAGGAGCAGGGGCTGATCAAGCCATCCAGGAACACTGGAGGACTTGTCCAGCCTTGAAGAAGCTCTAGTGTT
TCTGAATCTAGCCACTTGGCGGTAAGCATGATGCAACTTCTGCAACTTCTGCTGGGGCTTTTGGGGCCAGGTGG
CTACTTATTTCTTTTAGGGGATTGTGAGGAGGTGACCACTCTCACGGTGAAATACCAAGTGTGAGAGGAAGTGCC
ATCTGGTACAGTGATCGGGAAGCTGTCCAGGAAGTGGGCCGGGAGGAGAGCGGAGGCAAGCTGGGGCCGCCTT
CCAGGTGTTGAGCTGCTCAGGCGCTCCCAATTAGGTGGACTCTGAGGAAGGCTTGCTCAGCACAGGCAGGCG
GCTGGATCGAGAGCAGCTGTGCCGACAGTGGGATCCCTGCCTGGTTTCTTGATGTGCTTGCCACAGGGGATT
GGCTCTGATCCATGTGGAGATCCAAGTGTGGACATCAATGACCACAGCCAGGTTTCCCAAAGGCGAGCAGGA
GCTGGAATCTCTGAGAGCGCTCTCTGCGAACCCTGGATCCCTTGGACAGAGCTCTTGACCCAGACACAGGCCC
TAACACCCTGCACACCTACACTCTGTCTCCAGTGAGCACTTTGCCTTGGATGTGATTGTGGGCCCTGATGAGAC
CAACATCGAGAACTCATAGTGGTGAAGGAGCTGGACAGGGAATCCATTCAATTTTTTGATCTGGTGTAACTGC
CTATGACAATGGGAACCCCCAAGTCAGGTACCACTTGGTCAAGGTCAACGTCTTGAGCTCCAATGACAATAG
CCCTGCGTTTGTGAGAGTTCACTGGCACTGGAAATCCAAGAAGATGCTGCACCTGGTACGCTTCTCATAAACT
GACCGCCACAGACCTGACCAAGGCCCAATGGGGAGGTGGAGTTCTTCTCAGTAAGCACATGCCTCCAGAGGT
GCTGGACACCTTCACTATTGATGCCAAGACAGGCCAGGTCACTTCTGCGTCGACCTCTAGACTATGAAAAGAACC
TGCTACGAGGTGGATGTTCAAGCAAGGACCTGGGTCCCAATCCTATCCAGCCCATTTGCAAGTTCTCATCAA
GGTCTGAGTGTCAATGACAACATCCCAAGCATCCACGTCAACATGGGCTCCAGCCATCACTGGTGTGAGAAGC
TCTTCCCAAGGACAGTTTATTGCTCTTGTGATGGCAGATGACTTGGATTCAAGACACAATGGTTTGGTCCACTG
CTGGCTGAGCCAGAGCTGGGCCACTTCAGGCTGAAAAGAACTAATGGCAACACATACATGTTGCTAACCAATGC
CACACTGGACAGAGAGCAGTGGCCCAATATACCCTCACTCTGTTAGCCCAAGACCAAGGACTCCAGCCCTTATC
AGCCAAGAAACAGCTCAGCATTGATCAGTACATCAACGACAATGCACTGTGTTGAGAAAAGCAGGTATGA
AGTCTCCAGCGGGAACCACTTACCCTCTCTTCACTTACCTATTACCATCAAGGCTCATGATGAGCTGGGAT
TAATGAAAAGTCTCATACCGCATCCAGGACTCCCACTTGTCTCACTTAGTAGCTATTGACTCCAACACAGGAGA
GGTCACTGCTCAGAGGTCACTGAACTATGAAGAGATGGCCGGCTTTGAGTTCCAGGTGATCGCAGAGGACAGCGG
GCAACCCATGCTTGCATCCAGTGTCTGTGTGGGTGAGCCTCTTGGATGCCAATGATAATGCCCCAGAGGTGGT
CCAGCCTGTGCTGACGATGGAAGCCAGCCTCTCCGTGCTTGTGAATGCCTCCACAGGCCACCTGCTGGTGCC
CATCGAGACTCCCAATGGCTTGGGCCAGCGGGCACTGACACACCTCCACTGGCCACTCAGAGTCCCGGCCATT
CCTTTTGACAACCATTGTGGCAAGAGATGCAGACTCGGGGGCAATGGAGAGCCCCCTCTACAGCATCCGCAATGG
AAATGAAGCCACCTCTTCTATCTCAACCCCTCATACGGGGCAGCTGTTGCTCAATGTACCAATGCCAGCAGCCT
CATTGGGAGTGAGTGGGAGCTGGAGATAGTAGTAGAGGACCAGGGAAGCCCCCTTACAGACCCGAGCCCTGTT
GAGGTGATGTTTGTCAACAGTGTGGACCACTGAGGCACTCAGCCCGCAAGCCTGGGGCCTTGAGCATGTGAT
GCTGACGGTGATCTGCCTGGCTGACTGTTGGGCATCTTCCGGTTGATCTGGCTTGTTCATGTTCCATCTGCCG
GACAGAAAAGAGGACAACAGGGCCTACAAGTGTGGGAGGCGGAGTCCACCTACCGCCAGCAGCCCAAGAGGCC
CCAGAAACACATTGAGAAGGCAGACATCCACCTCGTGCCTGTGCTCAGGGGTGAGGAGGTGAGCCTTGTGAAGT
CGGGCAGTCCCAAAAGATGTGGACAAGGAGGCGATGATGGAAGCAGGCTGGGACCCCTGCCTGCAGGCCCTT
CCACCTCAGCCGACCTGTACAGGACGCTGCTAATCAAGGCAACAGGGAGCAGCGGCGAGAGCCGAGAGGT
GCTGCAAGACAGGTCAACCTCCTTTTCAACCATCCAGGCAGAGGAATGCCTCCCGGGAACCTGAACCTTCC
CGAGCCCCAGCCTGCCACAGGCCAGCCAGTTCAGGCCTCTGAAGGTTGCAGGCAGCCCCACAGGGAGGCTGGC
TGGAGACCAGGGCAGTGAGGAAGCCCCACAGAGGCCACAGCCTCCTCTGCAACCTGAGACGGCAGCGACATCT
CAATGGCAAGTGTCCCTGAGAAAGAAATCAGGGCCCCGTGATCCTGCGGAGCCTGGTCCGGCTGTCTGTGGC
TGCCTTCGCCGAGCGGAACCCGTGGAGGAGCTCACTGTGGATTCTCCTCTGTTGAGCAATCTCCAGCTGCT
GTCCTTGTGTCATCAGGGCAATTCCAGCCCAAAACCAACCAGGAGGAATAAGTACTTGGCCAAGCCAGGAGG
CAGCAGGAGTGCAATCCAGACACAGATGGCCCAAGTGAAGGGCTGGAGGCCAGACAGCCAGAACAGGAGGA
AGGGCCTTTGGATCCTGAAGAGGACCTCTCTGTGAAGCAACTGCTAGAAGAAGAGCTGTCAAGTCTGCTGGACCC
CAGCACAGGTCTGGCCCTGGACCGGCTGAGCGCCCTGACCCGGCCTGGATGGCGAGACTCTCTTGGCCCTCAC
CACCACCTACCGTGACAATGTGATCTCCCGGATGCTGACGCCAGGAGGAGCCGAGGACCTTCCAGACGTTCCG
CAAGGCAGAGGCACAGAGCTGAGCCCAACAGGCACGAGGCTGGCCAGCACCTTGTCTCGGAGATGAGCTCACT
GCTGGAGATGCTGCTGGAACAGCGCTCCAGCATGCCCGTGGAGGCGCCTCCGAGGCGCTGCGGCGGCTCTCGGT
CTGCGGGAGGACCTCAGTTTAGACTTGGCCACCAAGTGCAGCCTCAGGCATGAAAGTGAAGGGGACCCAGGTGG
AAAGACGGGGACTGAGGGCAAGAGCAGAGGCAGCAGCAGCAGCAGGTCCTGTAACATACCTCAGACGCCT
CTGGATCCAAGAACAGGGGCTGAGGATCTGTGACCAAGAGCTGGTTTCTAAAATCTTGAATCACTAGCTAG
CGGGCGCTGAGAATTTAGGGTGACTGATGCTAACCCCAAGAGGAGGCAAGAGCCCAAGGACTAACAGCTGAC
TGACCAAGCAGCCCTTGTAAAGCAGCTCTGAGTCTTTTGGAGGACAGGGACGGTTTGTGGCTGAGATAAGTGT
TCCTGGCAAAACATATGTGAGACACAAAGGCTCAGTCTCTGGCAGAACAGATGCCACGGAGTATCACAGGCAGG
AAAGGGTGGCCTTCTTGGGTAGCAGGAGTCAGGGGGCTGTACCCTGGGGGTGCCAGGAATGCTCTCTGACCTAT
CAATAAAGGAAAAGCAGTAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGG

394/550

FIGURE 394

MMQLLQLLLGLLPGGYLFLLGDCQEVTTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQAG
AAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFVLTATGDLALIHVEIQVL
DINDHQPRFPKGEQELEISESASLRTRIPLDRALDPDTGPNLTHTYTLSPSEHFALDVIVGPD
ETKHAELIVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSNDNSPFAESSLALE
IQEDAAPGTLLIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPLDYEKN
PAYEVDVQARDLGPNPIPAHCKVLIKVLDVNDNIPSIHVTWASQPSLVSEALPKDSFIALVMA
DDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQDQGLQPLS
AKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHDADLGINGKVSRYRIQDSPV
AHLVAIDSNTGEVTAQRSNLNVEEMAGFEFQVIAEDSGQPMLASSVSVWVSLLDANDNAPEVVQ
PVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPLATHSSRPFLTTIVARDADSGA
NGEPLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGSPLQTRALLRV
MFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEKKDNRAYNCREAE
STYRQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEAGWDPCLOAPFHLLT
PTLYRTLNRNQGNQGAPEAREVLQDTVNLLFNHPRQRNASRENLNLEPQPATGQPRSRPLKV
AGSPTGRLAGDQGSSEAPQRPASSATLRRQRHLNGKVSPEKESGPRQILRSLVRLSVAFAE
RNPVEELTVDSPPVQQISQLLSLLHQGFQPKPNHRGNKYLAKEGGSRSAPDTPDGPSARAGG
QTDPEQEEGPLDPEEDLSVKQLLEELSSLLDPSTGLALDRLSAPDPAWMARLSLPLTTNYRD
NVISPDAAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSLEMLLEQRSSMPVEAASE
ALRRLSVCGRTLSDLATSAASGMKVQGDPPGKTGTGEGKSRGSSSSSRCL

Important features:**Signal peptide:**

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

395/550

FIGURE 395

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAGG
CTCCAGAATCGTGTAACAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAATC
AGTAGGTGACCCCGCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCCGACCTCGTGCG
GCCAAGACGTGGATGTTCTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGCACAG
GAGGACAAGGTGCTGGGGGGTTCATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGGCCTTG
TTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAAGTGGCAACTGGGTCCCTTACAGCT
GCCCCACTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAATAAAGAT
GGCCCAGAGCAAGAAATACCTGTGGTTTCAGTCCATCCCACACCCCTGCTACAACAGCAGCGAT
GTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCCCTGGGGTCC
AAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGCACCGTCTCA
GGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCTGACACTCTCAACTGTGCAGAAGTA
AAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACAGATGGCATGGTC
TGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGGCCCCCTGGTGTGT
GATGGTGCCTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGAGGTCCGACAAACCT
GGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATCATAGGCAGCAAGGGC
TGATTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACTCTCTGGTTC

396/550

FIGURE 396

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVLV
GGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLLQL
RDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCEDAYP
GQITDGMVCAGSSKGADTCQGDSSGGLVCDGALQGITSWGSDPCGRSDKPGVYTNICRYLDWI
KKIIGSKG

Important Features:**Signal peptide:**

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

397/550

FIGURE 397

GGCGGCTGCTGAGCTGCCTTGAGGTGCAGTGTTGGGGATCCAGAGCCATGTCGGACCTGCTAC
TACTGGGCCTGATTGGGGGCCTGACTCTCTTACTGCTGCTGACGCTGCTGGCCTTTGCCGGGT
ACTCAGGGCTACTGGCTGGGGTGGAAGTGAGTGCTGGGTCACCCCCCATCCGCAACGTCACTG
TGGCCTACAAGTTCCACATGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGCTGCA
GCATCTCTCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCCCCCTG
ATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCCCTGAGC
TCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGCCATGTGG
TGACAGCCACCTTCCCCTACACCACCATTCTGTCCATCTGGCTGGCTACCCGCCGTGTCCATC
CTGCCTTGGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGGAGATCTACC
AGGAAGACCAGATCCATTTTCATGTGCCCACTGGCACGGCAGGGAGACTTCTATGTGCCTGAGA
TGAAGGAGACAGAGTGGAATGGCGGGGGCTTGTGGAGGCCATTGACACCCAGGTGGATGGCA
CAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGGAAGTGAGCCCTGGCAGCCGGG
AGACTTCAGCTGCCACACTGTCACCTGGGGCGAGCAGCCGTGGCTGGGATGACGGTGACACCC
GCAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGAGGAGCTGGACTTGG
AGGGCGAGGGGCCCTTAGGGGAGTCACGGCTGGACCCTGGGACTGAGCCCCTGGGGACTACCA
AGTGGCTCTGGGAGCCCACTGCCCCTGAGAAGGGCAAGGAGTAACCCATGGCCTGCACCCTCC
TGCAGTGCAGTTGCTGAGGAACTGAGCAGACTCTCCAGCAGACTCTCCAGCCCTCTTCCTCCT
TCCTCTGGGGGAGGAGGGGTTCTTGAGGGACCTGACTTCCCCTGCTCCAGGCCTCTTGCTAAG
CCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCAGGGACTATTTTCTGCACCA
GCCCCAGGGCTGCCGCCCTGTTGTGTCTTTTTTTTCAGACTCACAGTGAGCTTCCAGGACC
CAGAATAAAGCCAATGATTTACTTGTTTCACCTGGAAAAAAAAAAAAAAAAAAAA

398/550

FIGURE 398

MSDLLLLGLIGGLTLLLLLTLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGLYGETGRL
FTESCSISPKLRSIAVYYDNPHMVPPDKCRAVGSI LSEGEESPSPELIDLYQKFGFKVFSFP
APSHVVTATFPYTTILSIWLATRRVHPALDTYIKERKLCAYPRLEIYQEDQIHMCPLARQGD
FYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSAA TLSPGASSRGW
DDGDTRSEHSYSESGASGSSFEELDLEGEGLGESRLDPGTEPLGTTKWLWEPTAPEKGKE

400/550

FIGURE 400

MSNSVPLLCLFWSLCYCFAAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRTSKDPEHEG
CYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSALHTREKDANVVVVVDWLPL
AHQLYTDVNNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTVGRI
TGLDPAGPMFEGADIAHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIYPNGGDFQPGC
GLNDVLGSIAYGTITEVVKCEHERAVHLFVDSLQNQDKPSFAFQCTDSNRFKKGICLSCKNR
CNSIGYNAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP

Important features:**Signal peptide:**

amino acids 1-16

Lipases, serine active site.

amino acids 163-172

N-glycosylation sites.

amino acids 80-83 and 136-139

401/550

FIGURE 401

CTTCCCAGCCCTGTGCCCCAAAGCACCTGGAGCATATAGCCTTGCAGAACTTCTACTTGCCTG
CCTCCCTGCCTCTGGCCATGGCCTGCCGGTGCCTCAGCTTCCTTCTGATGGGGACCTTCCTGT
CAGTTTCCCAGACAGTCCTGGCCCAGCTGGATGCACTGCTGGTCTTCCCAGGCCAAGTGGCTC
AACTCTCCTGCACGCTCAGCCCCCAGCACGTACCATCAGGGACTACGGTGTGTCCTGGTACC
AGCAGCGGGCAGGCAGTGCCCCCTCGATATCTCCTCTACTACCGCTCGGAGGAGGATCACCACC
GGCCTGCTGACATCCCCGATCGATTCTCGGCAGCCAAGGATGAGGCCCAATGCCTGTGTCC
TCACCATTAGTCCCGTGCAGCCTGAAGACGACGCGGATTACTACTGCTCTGTTGGCTACGGCT
TTAGTCCCTAGGGGTGGGGTGTGAGATGGGTGCCTCCCCTCTGCCTCCCATTCTGCCCCCTGA
CCTTGGGTCCCTTTTAAACTTTCTCTGAGCCTTGCTTCCCCTCTGTAAAATGGGTTAATAATA
TTCAACATGTCAACAAC

402/550

FIGURE 402

MACRCLSFLMGTFLSVSQTVLAQLDALLVFPQVAQLSCTLSPQHVTIRDYGVSWYQQRAGS
APRYLLYYRSEEDHHRPADI PDRFSAKDEAHNACVLTISPVPEDDADYYCSVGYGFSP

403/550

FIGURE 403

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCGTGCAGCAGCTCCAGAAAGCAGCGAGTTG
GCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGTGGCTCACAACAAGATGCTCAAGGTGTCAGC
CGTACTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCAGTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGG
GCGGTGCGACGGCGGTAATTTTCTGGATGATAAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGG
ACAGTGAACAAATTCGAGACGAAGTAGAGGATGATTATTTCCGCACTTGGAGTCCAGGAAAACCTTTCGATCA
GGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTATGCATTGCTCAAGATTG
TCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATGAAAGAAGCAGGAGTAGACCATAGGCAGTG
GAGGGGTCCCATAATTATCCACCTGCAAGCAGTGCCAGTGGTCTATCCCAGCCCTGTTTGTGGTTCAGATGGTCA
TACCTACTCTTTTTCAGTGCAAACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGG
ACATTGCCCATGTCTTCAGATAAGCCACAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCTGGAGTT
CAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAGGCCCCTTCATGAAAGTGGAAGTCAAAACAAGAAGACAAA
AACATTGCTGAGGCCTGAGAGAAGCAGATTTCGATACCAGCATCTTGCCAATTTGCAAGGACTCACTTGGCTGGAT
GTTTAACAGACTTGATACAACTATGACCTGCTATTGGACCAGTCAGAGCTCAGAAGCATTACCTTGATAAGAA
TGAACAGTGTACCAAGGCATTCTTCAATTCTTGTGACACATACAAGGACAGTTTAATATCTAATAATGAGTGGTG
CTACTGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCAAGGGGTA
GAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGCCAACACAATGTGCATGGCAGTGT
TGGACAGTGTGGTGTGTTGACAGATATGGAAATGAAGTCATGGGATCCAGAATAAATGGTGTGAGATTGTGC
TATAGATTTTGAGATCTCCGGAGATTTTGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGA
CGATATTATGAATGATGAAGATGAAATGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGA
CCATGATGTATACATTTGATTGATGACAGTTGAAATCAATAAATCTACATTTCTAATATTTACAAAAATGATAG
CCTATTTAAATATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCACATATATTTGTATAATTATTTGAA
AAATTGCAGCTAAAGTTATAGAACTTTATGTTTAAATAAGAATCATTGCTTTGAGTTTTATATTCCTTACACA
AAAAGAAAATACATATGCAGTCTAGTCAGACAAAATAAAGTTTGAAGTGCTACTATAATAAATTTTCACGAGA
ACAACTTTGTAAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTGAAAGAT
AATTCTAAGTGAAATTTAAATAAATAAATTTTAATGACCTGGGTCTTAAGGATTTAGGAAAAATATGCATGCT
TTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAGGATAACAGAGAGATACCACATGACTCCA
AAAAAAAAAAAAA

404/550

FIGURE 404

MLKVSAVLCVCAAAWCSQSLAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFRD
EVEDDYFRTWSPGKPFQALDPAKDPCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEAGV
DHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCPSPDK
PTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKKTKTLLRPERSRFDTSILPICKDS
LGWMFNRLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQRQQDP
PCQTELSNIQKRQGVKKLLGQYIPLCDEDGYKPTQCHGSVGQCWCVDRYGNEVMGSRINGVA
DCAIDFEISGDFASGDFHEWTDDEDEDDIMNDEDEIEDDEDEDEGDDDDGGDDHDVYI

Important features:**Signal peptide:**

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 405

CGAAGGGGAGGAGCAGGCCACACAGGCACAGGCCGGTGAGGGACCTGCCACAGCTGGAGGGTCTCGCTCTGTCA
CACAGGCTGGAGTGCAGTGGTGTGATCTTGGCTCATCGTAACCTCCACCTCCCGGGTCAAGTGATTCTCATGCC
TCAGCCTCCCGAGTAGCTGGGATTACAGGTGGTGACTTCCAAGAGTACTCCGTGCGAGGAAAAATGACTCCCCAG
TCGCTGCTGCAGCAGCACTGTTCTTGCTGAGTCTGCTTCTGGTCCAAGGTGCCACGGCAGGGGCCACAGG
GAAGACTTTTCGCTTCTGCAGCCAGCGGAACAGACACAGAGGAGCGCTCCACTACAACCCACACAGACCTG
CGACTCTCCATCGAGAATCCGAAGAGGGCCCTCACAGTCCATGCCCCCTTTCCTGCAGCCCACCTGCTTCCCGA
TCCTTCCCTGACCCAGGGGCTCTACCACTTCTGCCTCTACTGGAACGACATGCTGGGAGATTACATCTTCTC
TATGGCAAGCGTGACTTCTTGCTGAGTGACAAAGCCTCTAGCCTCCTCTGCTTCCAGACCAGGAGGAGCGCTG
GCTCAGGGGCCCCCGCTGTTAGCCACTTCTGTCACTCTGCTGGAGCGCTCAGAACATCAGCTGCCAGTGGC
GCCAGGTTCACTTCTCTCTTCCACAGTCTCTCCCCACAGGGCGCTCACAATGCCTCGGTGGACATGTGCGAGCTC
AAAAGGGACCTCCAGCTGCTCAGCCAGTTCTCTGAAGCATCCCCAGAAGGCCTCAAGGAGGCCCTCGGCTGCCCCC
GCCAGCCAGCAGTTGCAGAGCCTGGAGTCGAAACTGACTCTGTGAGATTCAATGGGGACATGGTGTCTTCCAG
GAGGACCGGATCAACGCCACGGTGTGGAAGTCCAGCCCCACGGCCCTCCAGGACCTGCACATCCACTCCCGG
CAGGAGGAGGAGCAGAGCGAGATCATGAGTACTCGGTGCTGCTGCTCGAACACTTCTCAGAGGACGAAAGGC
CGGAGCGGGGAGGCTGAGAAGAGACTCTCTGCTGGTGGACTTCAGCAGCCAAGCCCTGTTCCAGGACAAGAATTCC
AGCCAAGTCTGGGTGAGAAGGTCTTGGGGATTGTGGTACAGAACACCAAAGTAGCCAACCTCACGGAGCCCCGTG
GTGCTCACTTTCCAGCACCAGCTACAGCCGAAGAATGTGACTCTGCAATGTGTGTTCTGGGTTGAAGACCCACA
TTGAGCAGCCCGGGGCACTTGAGCAGTGTGGGTGTGAGACCGTCAGGAGAAACCAACATCCTGCTTCTGC
AACCACTTGACCTACTTTGCAGTGTGATGGTCTCTCGGTGGAGGTGACGCGGTGCACAAGCACTACCTGAGC
CTCCTCTCCTACGTGGGCTGTGTGCTCTGCCCCGCGCTGCTTGTCAACATTGCCGCTACCTCTGCTCCAGG
GTGCCCTGCCGTGCAGGAGGAAACCTCGGGACTACACCATCAAGGTGCACATGAACCTGCTGCTGGCCGTCTTC
CTGCTGGACACGAGCTTCTGCTCAGCGAGCCGGTGGCCCTGACAGGCTCTGAGGCTGGCTGCCAGCCAGTGCC
ATCTCTCTGCACTTCTCCCTGCTCACTGCTTCTTCTGATGGGCTCAGGGGTACAACCTCTACCGACTCGTG
GTGGAGTCTTTGGCACCATTGTCTCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTT
CTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAACATATGCCCCATCATCTTGGCTGTGCATAGGACTCCAGG
GGCTCATCTACCTTCCATGTGCTGGATCCGGGACTCCCTGGTCAGTACATCACCACCTGGGCTCTTCAGG
CTGGTGTTTCTGTTCAACATGGCCATGCTAGCCACCATGGTGTGTCAGATCTCGGGCTGGCCCCACACCCAA
AAGTGTGCATGTGCTGACACTGCTGGGCTCAGCCTGTGCTTGGCTGCCCCGCGCTGGGCTTGATCTTCTCTCC
TTTGCTTCTGGCACCTTCCAGCTTGTGCTCTCTACCTTTTCAGCATCATCACTCCTTCCAAGGCTTCTCATC
TTCATCTGGTACTGTTCATGCGGCTGCAGGCCCGGGGTGGCCCTCCCCCTCTGAAGAGCACTCAGACAGCGCC
AGGCTCCCCCATCAGCTCGGGCAGCACCTCGTCCAGCCGCACTTAGCCCTCCAGCCACCTGCCCATGTGATGAAG
CAGAGATCGGCGCTCGTGCACAGTCCGTGTGGCCCCGAGGCCAGGCCAGCCGAGCCAGTCAAGCCGAGACT
TTGGAAGGCCCAACGACCATTGAGAGATGGGGCGTTGCCATGGTGGACGGACTCCCGGGCTGGGCTTTTGAATTG
GCCTTGGGGACTACTCGGCTCTCACTCAGCTCCACGGGACTCAGAAGTGCGCCGCATGCTGCTAGGGTACTG
TCCCCACATCTGTCCCAACCCAGCTGGAGGCGCTGGTCTCTCTTACAACCCCTGGGCGCCAGCCCTCATTGCTGG
GGCCAGGCGCTTGGATCTTGAGGCTGTGGACATCTTAATCTGTGCCCTGCTGGGACGAAATGTGGCTCCA
GTTGCTCTGTCTCTGTGTTCACTTCACTGAGGCACTCTGATCTCTGTCAATTTAACCTCAGGTGGCACCCAGGG
CGAATGGGGCCCAGGGCAGACCTTCAGGGCCAGAGCCCTGGCGGAGGAGAGGCCCTTTGCCAGGAGCACAGCAGC
AGCTCGCCTACCTCTGAGCCAGGCCCCCTCCCTCCCTCAGCCCCCAGTCTCCCTCCATCTTCCCTGGGGTTC
TCCTCTCTCCCAGGGCCTCTTGTCTCTTCTGTTACAGTCTGGGGTCCCCGATTTCAATGCTGTTTTTGGGA
GTGGTTTCCAGGAGCTGCTGGTGTCTGCTGTAAATGTTGTCTACTGCAACAGCCTCGGCTGCCCCGAGCCA
GGCTCGGTACCGATGCGTGGGCTGGGCTAGGTCCCTCTGTCCATCTGGGCCCTTGTATGAGCTGCATTGCCCTTG
CTCACCTGACCAAGCACACGCCTCAGAGGGGCCCTCAGCTCTCTCTGAAGCCCTCTTGTGGCAAGAAGTGTGGA
CCATGCCAGTCCCGTCTGGTTTCCATCCCACCACTCCAAGGACTGAGACTGACCTCCTCTGGTGACACTGGCCTA
GAGCCTGACACTCTCTAAGAGTTCTCTCAAGCCCCAAATAGCTCCAGCGCCCTCGGCCGCCATCATGTT
TAATTCCTGTTCAACAACACACAGCGGTAGATTGCTGGCTGTTGTAGGTGGTAGGGACACAGATGACCGACCTG
GTCACTCTCTGCCAACATTGAGTCTGGTATGTGAGGCGTGGTGAAGCAAGAAGTCTTGGAGCTACAGGGACA
GGGAGCCATCATTTCTGCTGGGAATCTTGAAGACTTCTGCAAGGATCAGCGTTCAATCTTGACCTTGAAGAT
GGGAAGGATGTTCTTTTACGTACCAATTCTTTTGTCTTTTGATATTAAGAAAGTACATGTTCTATTGTAGAGA
ATTTGGAAGCTGTAGAAGAAATCAAGAAAAAATAAAATCAGCTGTTGTAATCGCTGCACAAAAAATAAAAAA
AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA

406/550

FIGURE 406

MT PQSLLQTTLFLLSLLFLVQGAHGRGHREDFRFCSQRNQTHRSSLHYKPTPDLRISIENSEE
ALT VHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLLSDKASSLLCFQHQE
ESLAQGPPLLATSVTSWWSPQNISLPSAASF TFSFHSPHTAAHNASVDMCELKRD LQLLSQF
LKHPQKASRRPSAAPASQQLQSLESKLT SVRFMGDMVSFEEDRINATVWKLOPTAGLQDLHIH
SRQEEEQSEIMEYSVLLPRTL FQRTKGRSGEAEKRLLLVD FSSQALFQDKNSSQVLGEKVLGI
VVQNTKVANLTEPVVLT FQHQLQPKNVT LQCVFWEDPTLSSPGHWSSAGCETVRRETQTSCF
CNHLTYFAVLMVSSVEVD AVHKHYLSLLSYVGCVV SALACLV TIAAYLCSRVP LPCRKRPRDY
TIKVHMNLLLA VFLDTSFLLSEPVALT GSEAGCRASAI FLHFSLLTCLSWMGLEGYNLYRLV
VEVFGTYVPGYLLKLSAMGWGFPI FLVTLVALVDVDNYGPI ILAVHRTPEGVIYPSMCWIRDS
LVSYITNLGLFSLVFLFN MAMLATMVVQILRLRPHTQKWSHVL TLLGLSLVLGLPWALIFFSF
ASGTFQLVVLYLFSIITSFQGFLIFIWYWSMRLQARGGPSPLKSNSDSARLPISSGSTSSSRI

Important features:**Signal peptide:**

amino acids 1-25

Putative transmembrane domains:amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657**Microbodies C-terminal targeting signal.**

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344**G-protein coupled receptors family 2 proteins**

amino acids 475-504

407/550

FIGURE 407

TTGTGACTAAAAGCTGGCCTAGCAGGCCAGGGAGTGCAGCTGCAGGCGTGGGGGTGGCAGGAG
CCGCAGAGCCAGAGCAGACAGCCGAGAAACAGGTGGACAGTGTGAAAGAACCAGTGGTCTCGC
TCTGTTGCCCAGGCTAGAGTGACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCTGGA
CTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACTCAAGCCC
TGTTTCTTCTCCTTCTGTGAGTGGACCACGGAGGCTGGTGAGCTGCCTGTCATCCCAAAGCTC
AGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGGCATAGAAGCCAGGAGCAGGGC
TCTCAGAAGGCGGTGGTGCCAGCTGGGATCATGTTGTTGGCCCTGGTCTGTCTGCTCAGCTG
CCTGCTACCCCTCCAGTGAGGCCAAGCTCTACGGTCGTTGTGAACTGGCCAGAGTGCTACATGA
CTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGGTCTGCCTTGCTTATTTAC
AAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACAACGGGATCTT
CCAGATCAACAGCCGGAGGTGGTGCAGCAACCTCACCCCGAACGTCCCCAACGTGTGCCGGAT
GTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGCCATGAAGATAAC
CCAAGAGCCTCAGGGTCTGGGTACTGGGAGGCCTGGAGGCATCACTGCCAGGGAAAAGACCT
CACTGAATGGGTGGATGGCTGTGACTTCTTAGGATGGACGGAACCATGCACAGCAGGCTGGGAA
ATGTGGTTTGGTTCCTGACCTAGGCTTGGAAGACAAGCCAGCGAATAAAGGATGGTTGAACG
TGAAA

408/550

FIGURE 408

MLLALVCLLSCLLPSSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALDY
EADGSTNNGIFQINSRRWCSNLTPNVPNVCRMYSDDLNPNLKDTVICAMKITQEPQGLGYWE
AWRHHCQGKDLTEWVDGCDF

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homologous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

409/550

FIGURE 409

CAGACTCCAGATTTCCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGACGC
CTCTTTTCAGCCCGGGATCGCCCCAGCAGGGATGCGGCGACAAGATCTGGCTGCCCTTCCCCGTGCTCCTTCTGGCC
GCTCTGCCTCCGGTGCTGCTGCCCTGGGGCGGCCGGCTTCACACCTTCCCTCGATAGCGACTTCACCTTTACCCCTT
CCCGCCGGCCAGAAGGAGTGCTTCTACCAGCCCATGCCCTGAAGGCCTCGCTGGAGATCGAGTACCAAGTTTTA
GATGGAGCAGGATTAGATATTGATTTCCATCTTGCTCTCCAGAAGGCAAAACCTTAGTTTTTGAACAAAGAAAA
TCAGATGGAGTTCACACTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAATACATTACGACCATT
TCTGAGAAGGTGATTTTCTTTGAATTAATCCTGGATAATATGGGAGAACAGGCACAAGAACAAGAAGATTGGAAG
AAATATATTACTGGCACAGATATATTGGATATGAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCC
AGACTAAGCAAAAGTGGGCACATACAAATTCTGCTTAGAGCATTGAAAGCTCGTGATCGAAACATACAAGAAAGC
AACTTTGATAGAGTCAATTTCTGGTCTATGGTTAATTTAGTGGTCATGGTGGTGGTGTGAGCCATTCAAGTTTAT
ATGCTGAAGAGTCTGTTTGAAGATAAGAGGAAAAGTAGAACTTAAACTCCAAACTAGAGTACGTAACATTGAAA
AATGAGGCATAAAATGCAATAAAGTGTACAGTCAAGACCATTAAATGGTCTTCTCCAAAATATTTTGAGATATA
AAAGTAGGAAACAGGTATAATTTAATGTGAAAATTAAGTCTTCACTTCTGTGCAAGTAATCCTGCTGATCCAG
TTGTACTTAAGTGTGTAACAGGAATATTTTGCAGAATATAGGTTTAACTGAATGAAGCCATATTAATACTGCAT
TTTCCTAACTTTGAAAAATTTTGCAAATGTCTTAGGTGATTTAAATAAATGAGTATTGGGCCCTAATTGCAACACC
AGTCTGTTTTTAAACAGGTTCTATTACCCAGAACCTTTTTGTAAATGCGGCAGTTACAAATTAAGTGTGGAGTTT
TCAGTTTTAAGTTATAAATCACCTGAGAATTACCTAATGATGGATTGAATAAATCTTTAGACTACAAAAGCCCAA
CTTTTCTCTATTACATATGCATCTCTCTATAATGTAATAGAATAATAGCTTTGAAATACAATTAGGTTTTTG
AGATTTTTATAACCAAAATACATTTCAAGTGAACATATTAGCAGAAAGCATTAGTCTTTGTACTTTGCTTACATTC
CCAAAAGCTGACATTTTACGATTCTTAAAAACACAAAGTTACACTTACTAAAATTAGGACATGTTTTCTCTTTG
AAATGAAGAATATAGTTTAAAGCTTCTCTCCATAGGGACACATTTTCTCTAACCTTAAGTAAAGTGTAGGA
TTTTAAAATTAATGTGAGGTAAATAAGTTTATTTTAAATAGTATCTGTCAAGTTAATATCTGTCAACAGTTAA
TAATCATGTTATGTTAATTTTAAACATGATTGCTGACTTGGATAATTCATTATTACCAGCAGTTATGAAGGAAATA
TTGCTAAAATGATCTGGGCCTACCATAAATAAATATCTCCTTTTCTGAGCTCTAAGAATTATCAGAAAACAGGAA
AGAATTTAGAAAACCTTGAAGAAACCTAATCCAAAATAAAATTCACTTAAGTAGAATATAAATAAATATCTAGA
ATCTGACTGGCTCATCATGACATCCTACTCATAACATAAATCAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG
AACTTTGGCTGTAGGTTTTTATTTTCTACAAGAATCTGGTTTGAATTATTTTGTAGCAGGTACATTTTATA
AAATGAAGCCCTACTGTAAGGTTTAGCACTGGGTGTACATATTTTATTAATAAATTTTATTTATATAACACTTTTAT
TAAAATGGCCTTTTCTGAACCTTTTATTTATGATGTTGAAGTAAGGATTAGAAACATAGACTCCCAAGTTTAAAA
CACCTAAATGTGAATAACCCATATATACAACAAAGTTTCTGCCATCTAGCTTTTTGAAGTCTATGGGGGTCTTAC
TCAAGTACTAGTAATTTAACTTCATCATGAATGAACATAATTTTAAAGTTATGCCCATTTATAACGTTGTTTAT
GACTACATTGTGAGTTAGAAACAACTTAAATTTGGGGTATAGAACCCTCAACAGGTTAGTAATGCTGGAATT
CTTGATGAGCAATAATGATAACCAGAGAGTGATTTTCACTTACACTCATAGTAGTATAAAAAGAGATACATTTCCC
TCTTAGGCCCTGGGAGAGAGCAGCTTAGATTTCCCTACTGGCAAGGTTTTTAAAAATGAGGTAAATGCCGTAT
ATGATCAATTACCTTAATTTGGCCAAGAAAATGCTTCAGGTGTCTAGGGGTATCCTCTGCACACTTGCAGAACAA
AGGTCAATAAGATCCTTGCCATGAATACCCCTCCCTTTTGGCTGTAAATTTGCAATGAGAAGCAAATTTACA
GTACCATAACTAATAAAGCAGGGTACAGATATAAACTACTGCATCTTTTCTATAAACTGTGATTAAGAATTCTA
CCTCTCCTGTATGGCTGTTACTGTACTGTACTCTCTGACTCCTTACCTAACATGAATTTGTTACATAATCTTCT
ACATGTATGATTTGTGCCACTGATCTTAAACCTATGATTCAGTAACTTCTTACCATATAAAAACGATAATTGCTT
TATTTGGAAGAATTTAGGAATACTAAGGACAATTTATTTTATAGACAAAGTAAAAAGACAGATATTTAAGAGG
CATAACCAAAAAGCAAACTTGTAAACAGAGTAAAAATCTTAAATATTCTAAAGACATACTGTTTATCTGCTT
CATATGCTTTTTTTAATTTCACTATTCCATTTCTAAATTAAGTTATGCTAAATTGAGTAAAGCTGTTTATCACTT
AACAGCTCATTTTGTCTTTTCAATATACAAATTTTAAAAATACTACAATATTTAACTAAGGCCCAACCGATTTT
CATAATGTAGCAGTTACCGTGTTACCTCACACTAAGGCCTAGAGTTTGCTCTGATATGCATTTGGATGATTAAT
GTTATGCTGTTCTTTTATGTGAATGTCAAGACATGGAGGGTGTGTAATTTTATGGTAAAAATTAATCCTTCTTA
CACATAATGGTGTCTTAAATTTGACAAAAATGAGCACTTACAATTGTATGCTCCTCAAATGAAGATTTCTTAT
GTGAAATTTTAAAGACATTGATTCGCGCATGTAAGGATTTTTCATCTGAAGTACAATAATGCACAATCAGTGTTG
CTCAAAGCTTTTATCTTATAAACAGCCATCTTAAATAAGCAACGTATTGTGAGTACTGATATGTATATAATAA
AAATTATCAAAGGAAAA

410/550

FIGURE 410

MGDKIWLPFPVLLLAALPPVLLPGAAGFTPSLSDFTFTLPAGQKECFYQPMPLKASLEIEYQ
VLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTFSTISEKVIFFELIL
DNMGEQAQEQEDWKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRNIQE
SNFDRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

411/550

FIGURE 411

CCCAGCTGAGGAGCCCTGCTCAAGACACGGTCACTGGATCTGAGAACTTCCCAGGGGACCGCATTCCAGAGTCA
GTGACTCTGTGAAGCACCCACATCTACCTCTTGCCACGTTCCACAGGGCTTGGGGGAAAGATGGTGGGGACCAAG
GCCTGGGTGTTCTCCTTCTGGTCCCTGGAAGTCACATCTGTGTTGGGGAGACAGACGATGCTCACCCAGTCAGTA
AGAAGAGTCCAGCCTGGGAAGAAGAACCCAGCATCTTTGCCAAGCCTGCCGACACCTGGAGAGCCCTGGTGAG
TGGACAACATGGTTCAACATCGACTACCCAGGCGGGAAGGGCGACTATGAGCGGCTGGAGGCCATTGCTTCTAC
TATGGGGACCGTGTATGTGCCGCTCCCTGCGGCTAGAGGCTCGGACCACTGACTGGACACCTGCGGGGAGCACT
GGCCAGGTGGTCCATGGTAGTCCCCGTGAGGGTTTCTGGTGCCTCAACAGGGGAGCAGCGGCTGGCCAGAAGTGC
TCTAATTACACCGTAGCTTCCCTGCCCCACAGGATCCCTGCGCCGAGACACAGAGCGCATCTGGAGCCCATGG
TCTCCCTGGAGCAAGTGTCTAGCTGCCTGTGGTCAAGTGGGGTCCAGACTCGCACACGATTGCTTGGCAGAG
ATGGTGTGCTGTGAGTGGAGCCAGGGAAGAGGGTCAAGTGCATGGGCCAGGACTGTACAGCCTGTGACCTG
ACCTGCCCAATGGGGCAGGTGAATGCTGACTGTGATGCCTGCATGTGCCAGGACTTCATGCTTCATGGGGCTGTC
TCCCTTCCCAGGAGTCCCCAGCCTCAGGGGCTGCTATCTACCTCCTGACCAAGACGCCGAAGCTGCTGACCCAG
ACAGACAGTGTGGGAGATTCCGAATCCCTGGCTTGTGCCCTGAAGGCAGCCACCATCAAGGCAGAGTTTGTGAGG
AAGTTTGGCCCCATGTACTCACAAATGCCCAAGACTAGCCCTGAAGGCAGCCACCATCAAGGCAGAGTTTGTGAGG
GCAGAGACTCCATACATGGTGTGAACCCCTGAGACAAAAGCAGGAGAGCTGGGCAGAGCGTGTCTCTGTGCTGT
AAGGCCACAGGGAAGCCAGGCCAGACAAGTATTTTGGTATCATAATGACACATTGCTGGATCCTTCCCTCTAC
AAGCATGAGAGCAAGCTGGTGTGAGGAACTGCAGCAGCACCAGGCTGGGGAGTACTTTTGAAGGCCACAGAGT
GATGCTGGGGCTGTGAAGTCCAAGGTTGCCAGCTGATTGTACAGCATCTGATGAGACTCCTTGAACCCAGT
CCTGAGAGCTATCTTATCCGGCTGCCCATGATTGCTTTGAGAATGCCACCAACTCCTTCTACTATGACCTGGGA
CGCTGCCCTGTTAAGACTTGTGCAGGGCAGCAGGATAATGGGATCAGGTGCCGTGATGCTGTGCAGAAGTGTGT
GGCATCTCAAGACAGAGGAAAGGGAGATCCAGTGCAGTGGCTACACGCTACCCACCAAGGTGGCCAGGAGTGC
AGCTGCCAGCGGTGTACGGAAACTCGGAGCATCGTGGGGCCGTGTGCTGCTGACAATGGGGAGCCCATG
CGCTTGGCCATGTGTACATGGGGAACAGCCGTGTAAGCATGACTGGCTACAAGGGCACTTTACCCCTCATGTG
CCCCAGGACACTGAGAGGCTGGTGTCTACATTTTGGACAGGCTGCAGAAGTTTGTCAACACCAAGTGTGCTA
CCTTTCAACAAGAAGGGGAGTGCCGTGTTCCATGAATCAAGATGCTTCCGCGAAAGAGCCCATCACTTTGGAA
GCCATGGAGACCAACATCATCCCCCTGGGGGAAGTGGTGGTGAAGACCCCATGGCTGAACTGGAGATTCCATC
AGGAGTTTCTACAGGCAGAATGGGGAGCCCTACATAGGAAAGTGAAGGCCAGTGTGACCTTCTGGATCCCTGG
AATATTTCCACAGCCACAGCTGCCAGACTGACCTGAACCTTCATCAATGACGAAGGAGACACTTTCCCCCTCGG
ACGTATGGCATGTTCTCTGTGGACTTCAGAGATGAGGTCACTCAGAGCCACTTAATGCTGGCAAAGTGAAGGTC
CACTTTGACTCGACCCAGGTCAAGATGCCAGAGCAGATATCCACAGTGAACCTCTGGTCACTCAATCCAGACACA
GGGCTGTGGGAGGAGGAGGTGATTTCAAATTTGAAAAATCAAAGGAGGAACAAAAGAGAAGACAGAACCTTCTG
GTGGGCAACCTGGAGATTCTGTGAGAGGAGGCTTTTAACTGGATGTTCTGAAAGCAGGCGGTCTTTGTAAAG
GTGAGGGCTACCGGAGTGAGAGGTTCTTGCTTGTGAGCAGATCCAGGGGGTGTGATCTCCGTGATTAACTG
GAGCCTAGAAGTGGCTTCTGTCCAACCTAGGGCCTGGGGCCGCTTTGACAGTGTCTACAGGCCCCAACGGG
GCCTGTGCTGCCCTTCTGTGATGACCACTCCCTGTGCTTACTCTGCTATGCTTGGCAAGCCTTGGCTGGG
GAGGAAGTCAAGCAGTGGAGTCTTCTCTAAATTCAACCCAAATGCAATTGGCGTCCCTCAGCCCTATCTCAAC
AAGCTCAACTACCGTCGGACGGACCATGAGGATCCACGGGTTAAAAAGACAGCTTTCCAGATTAGCATGGCCAG
CCAAGGCCCAACTCAGCTGAGGAGAGCAATGGGCCCATCTATGCCTTTGAGAACCTCCGGGCATGTGAAGAGGCA
CCACCCAGTGCAGCCCACTTCCGGTCTACAGATTGAGGGGGATCGATATGACTACAACACAGTCCCCCTCAAC
GAAGATGACCCATGAGCTGGACTGAAGACTATCTGGCATGGTGGCCAAAGCCGATGGAATTCAGGGCCTGCTAT
ATCAAGGTGAAGATTGTGGGGCCACTGGAAGTGAATGTGCGATCCCGCAACATGGGGGCACTCATCGGCGGACA
GTGGGGAGCTGTATGGAATCCGAGATGTGAGGAGCACTCGGGACAGGGACAGCCCAATGTCTCAGCTGCCGTG
CTGGAGTTCAAGTGCAGTGGGATGCTCTATGATCAGGACCGTGTGGACCGCACCCCTGGTGAAGGTCACTCCCGAG
GGCAGCTGCCGTGAGCCAGTGTGAACCCCATGCTGCATGAGTACCTGGTCAACCACTTGCCACTTGCACTCAAC
AACGACACCAAGTGAATACCATGCTGGCACCTTGGACCCACTGGGGCACAACATATGGCATCTACACTGTCACT
GACCAGGACCTCGCACGGCCAAGGAGATCGCGCTCGGGCGGTGCTTTGATGGCACATCCGATGGCTCCTCCAGA
ATCATGAAGAGCAATGTGGGAGTAGCCCTCACCTTCAACTGTGTAGAGAGGCAAGTAGGCCGCCAGAGTGCCTTC
CAGTACCTCAAAGCACCCAGCCAGTCCCTGCTGCAGGCACTGTCCAAGGAAGAGTGCCTCGAGGAGGCAG
CAGCGAGCGAGCAGGGGTGGCCAGCGCCAGGGTGGAGTGGTGGCCTCTCTGAGATTCTCTAGAGTTGCTCAACAG
CCCCTGATCAACTAAGTTTGTGGTACTTCACTCTCTGCTGCTCATTTCATGTGACAGCCATTGTGAGACTGA
TGCACAACTGTCACTTGGTTAATTTAAGCACTTCTGTTTTCGTGAATTTGCTTGTGTTTCTTCTCATGCCCTTA
CTTACTTTGTCCCATGCTACTGATTGGCAGTGGCCCCACAATGGCACAATAAAGCCCCCTTGTGAACTGTTC
TTTAAATGAAACACAAGAAATTTGGCCACTGGTAAACTCTGCAGCTTCAACTGTACTTCAATTAATGCCATTAAT
GCAATATACTTCTCTTTTTCATGGTTTGGCCACCTCTGCAATAGTGATAATCTGATGCTGAAGATCAA
ATAACCAATATAAAGCATATTTCTTGGCCTTGTCTCCACAGGACATAGGCAAGCCTTGATCATAGTTCATACATAT
AAATGGTGGTGAATATAAGAAATAAACACAATACTTTTACTTGAATGTAAATAACTTATTTATTTCTTTGCTA
AATTTGGAATTCTAGTGACATTCAAAGTTAAGCTATTAAATATAGGGTGATCATAGTTCCCTACCAAGTCTGG
AAAGAACATCTCTGGTATCCACAATTACACAGGTTGCTAACTGTATTTGTACATTTCCCTTGTGATTGCGTTT
TGTCTTGTCTAGAAACCCAGTGTAGCCAGGGCAGATGTCAATAATGCATACTCTGTATTTGAAAAA

412/550

FIGURE 412

MVGTKAWVFSFLVLEVTSVLGRQTM LTQSVRRVQPGKKNPSIFAKPADTLESPGEWTTWFNID
YPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGQVVHGSPREGFWCLNREQRP
GQNC SNYTVRFLCPPGSLRRDTERI WSPWSPWSKCSAACGQTGVQTRTRICLAEMVSLCSEAS
EEGQH CMGQDCTACDLTC PMGQVNADCDACMCQDFMLHGAVSLPGGAPASGAAIYLLTKTPKL
LTQTDSDGRFRIPGLCPDGKSILKITKVKFAPIVLTMPKTS LKAATIKAEFVRAETPYMVMNP
ETKARRAGQSVSLCCKATGKPRPD KYFWYHNDTLLDPSLYKHESKLVLRLKQQHQAGEYFCKA
QSDAGAVKSKVAQLIVTASDETPCNPVPESYLIRLPHDCFQ NATNSFYDVGRCPVKTCAGQQ
DNGIRCRDAVQNC CGISKTEEREIQCSGYTLPTKVAKECSCQRC TETR SIVRGRVSAADNGEP
MRFGHVYMGNSRVSM TGYKGTFTLHVPQDTERLVLT FVDRLQKFVNTTKVL PPNKKGS AVFHE
IKMLRRKEPIT LEAMETNIIPLGEVVGEDPMAELEIPSR SFYRQNGEPYIGKVKASVTFLDPR
NISTATAAQTD LNF INDEGDTFPLRTYGMFSVDFRDEVTSEPLNAGKVKVHLDSTQVKMPEHI
STVKLWSLNPDTGLWEEEGDFKFENQRRNKREDRTFLVGNLEIRERRLE NLDVPESRRCFVKV
RAYRSE RFLPSEQIQGVVISVINLEPRTGFLSNPRAWGRFDSVITGPN GACVPAFCDDQSPDA
YSAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLN YRRTDHEDPRVKKTAFQISMAKPR
PNSAEESNGPIYAFENLRACEEAPP SAAHFRFYQIEGDRYDYNTVPFNEDDPMSWTE DYLAWW
PKPMEFRACYIKVKIVGPLEVNVRSRNMGGTHRRTVGKLYGIRDVRSTRDRDQPNVSAACLEF
KCSGMLYDQDRVDRTL VKVIPQGS CRRASVNPMLHEYLVNHLPLAVNNDTSEY TMLAPLDPLG
HNYGIYTVTDQDPRTAKEIALGRCFDGTSDGSSRIMKSNVGVALTFNCVERQVGRQSAFQYLQ
STPAQSPAAGTVQGRVPSRRQQRASRGGQRQGGVVASLRFPRVAQQPLIN

413/550

FIGURE 413

GCCACGTTGTCTTCTTTCTTCACCACCACCCAGGAGCTCAGAGATCTAAGCTGCTTTCCATC
TTTTCTCCCAGCCCCAGGACACTGACTCTGTACAGGATGGGGCCGTCTTGCCTCCTTCTC
ATCCTAATCCCCCTTCTCCAGCTGATCAACCCGGGGAGTACTCAGTGTTCTTAGACTCCGTT
ATGGATAAGAAGATCAAGGATGTTCTCAACAGTCTAGAGTACAGTCCCTCTCCTATAAGCAAG
AAGCTCTCGTGTGCTAGTGTCAAAAGCCAAGGCAGACCGTCCTCCTGCCCTGCTGGGATGGCT
GTCACTGGCTGTGCTTGTGGCTATGGCTGTGGTTCGTGGGATGTTTCACTGGAAACCACTGC
CACTGCCAGTGCAGTGTGGTGGACTGGACCACTGCCCGCTGCTGCCACCTGACCTGACAGGGA
GGAGGCTGAGAACTCAGTTTGTGACCATGACAGTAATGAAACCAGGGTCCCAACCAAGAAAT
CTAACTCAAACGTCCCACTTCATTTGTTCCATTCTTGATTCTTGGGTAATAAAGACAACTTT
GTACCTCAAAAAAAAAAAAAAAAAAAAAA

414/550

FIGURE 414

MGPSSCLLLILIPLLQLINPGSTQCSLDSVMDKKIKDVLNSLEYSPSPISKKLSCASVKS

QGRPSSCPAGMAVTGCACGYGCGSWDVQLETTCHCQCSVVDWTTARCCHLT

415/550

FIGURE 415

CAGAAGAGGGGGCTAGCTAGCTGTCTCTGCGGACCAGGGAGACCCCCGCGCCCCCCCCGGTGTG
AGGCGGCCCTCACAGGGCCGGGTGGGCTGGCGAGCCGACGCGGCGGCGGAGGAGGCTGTGAGGA
GTGTGTGGAACAGGACCCGGGACAGAGGAACCAATGGGCTCCGCAGAACCTGAGCACCTTTTGCC
TGTTGCTGCTATACCTCATCGGGGCGGTGATTGCCGACGAGATTTCTATAAGATCTTGGGGG
TGCCCTCGAAGTGCCTCTATAAAGGATATTTAAAAGGCCTATAGGAACTAGCCCTGCAGCTTC
ATCCCGACCGGAACCCTGATGATCCACAAGCCCAGGAGAAATTCCAGGATCTGGGTGCTGCTT
ATGAGGTTCTGTCAGATAGTGAGAAACGGAAACAGTACGATACTTATGGTGAAGAAGGATTAA
AAGATGGTCATCAGAGCTCCCATGGAGACATTTTTTTCACACTTCTTTGGGGATTTTGGTTTCA
TGTTTGGAGGAACCCCTCGTCAGCAAGACAGAAATATTCCAAGAGGAAGTGATATTATTGTAG
ATCTAGAAGTCACTTTGGAAGAAGTATATGCAGGAAATTTTGTGGAAGTAGTTAGAAACAAAC
CTGTGGCAAGGCAGGCTCCTGGCAAACGGAAGTGCAATTGTGCGCAAGAGATGCGGACCACCC
AGCTGGGCCCTGGGCGCTTCCAAATGACCCAGGAGGTGGTCTGCGACGAATGCCCTAATGTCA
AACTAGTGAATGAAGAACGAACGCTGGAAGTAGAAATAGAGCCTGGGGTGAGAGACGGCATGG
AGTACCCCTTTATTGGAGAAGGTGAGCCTCACGTGGATGGGGAGCCTGGAGATTTACGGTTCC
GAATCAAAGTTGTCAAGCACCCAATATTTGAAAGGAGAGGAGATGATTTGTACACAAATGTGA
CAATCTCATTAGTTGAGTCACTGGTTGGCTTTGAGATGGATATTACTCACTTGGATGGTCACA
AGGTACATATTTCCCGGGATAAGATCACCAGGCCAGGAGCGAAGCTATGGAAGAAAGGGGAAG
GGCTCCCCAACTTTGACAACAACAATATCAAGGGCTCTTTGATAATCACTTTTGATGTGGATT
TTCCAAAAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAAACAGCTACTGAAACAAGGGT
CAGTGCAGAAGGTATACAATGGACTGCAAGGATATTGAGAGTGAATAAAATTGGACTTTGTTT
AAAATAAGTGAATAAGCGATATTTATTATCTGCAAGGTTTTTTTTGTGTGTGTTTTTGTTTTTA
TTTTCAATATGCAAGTTAGGCTTAATTTTTTTTATCTAATGATCATCATGAAATGAATAAGAGG
GCTTAAGAATTTGTCCATTTGCATTTCGAAAAGAATGACCAGCAAAAGGTTTACTAATACCTC
TCCCTTTGGGGATTTAATGTCTGGTGCTGCCGCCTGAGTTTCAAGAATTAAAGCTGCAAGAGG
ACTCCAGGAGCAAAAGAAACACAATATAGAGGGTTGGAGTTGTTAGCAATTTCAATCAAATG
CCAACTGGAGAAGTCTGTTTTTAAATACATTTTGTGTTATTTTTTA

416/550

FIGURE 416

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKALQLHPDRNPDDPQAQEKFQDLGAAYE
VLSDSEKRKQYDTYGEGLKDGHQSSHGDI FSHFFGDFGFMFGGT PRQQDRNIPRGSDIIVDLEVTLEEVYAGNF
VEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQMTQEVVVCDECPNVKLVNEERTLEVEIEPGVRDGM EYPFI
GEGEPHVDGEPGDLRFRIKVVKHPIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVVHISRDKITRPGAKLW
KKGEGLPNFDN NNIKGS LIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

Important features:**Signal peptide:**

amino acids 1-22

Cell attachment sequence.

amino acids 254-257

Nt-dnaJ domain signature.

amino acids 67-87

Homologous region to Nt-dnaJ domain proteins.

amino acids 26-58

N-glycosylation site.

amino acids 5-9, 261-265

Tyrosine kinase phosphorylation site.

amino acids 253-260

N-myristoylation site.

amino acids 18-24, 31-37, 93-99, 215-221

Amidation site.

amino acids 164-168

417/550

FIGURE 417

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCCAGCAGGA
TGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCCGGAGAGGGGCCAGCCCCGGGGCAG
GATGACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCGGTGTTTCATGATCCTGCT
GATCATCGTGTACTIONGGACAGCGCAGGCGCCGCGCACTTCTACTTGCACACGTCCTTCTCTAG
GCCGCACACGGGGCCCGCTGCCACGCCCCGGGCCGACAGGGACAGGGAGCTCACGGCCGA
CTCCGATGTGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGACCTTCC
CAGAAAGGAGACGGAGCAGCCGCCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAGGCTACGA
CTGGTCCCCGCGCGACGCCCGGCGCAGCCAGACCAGGGCCGGCAGCAGGCGGAGCGGAGGAG
CGTGCTGCGGGGCTTCTGCGCCAACCTCCAGCCTGGCCTTCCCCACCAAGGAGCGCGCATTCTGA
CGACATCCCCAAGCTCGGAGCTGAGCCACCTGATCGTGACGACCGGCACGGGGCCATCTACTG
CTACGTGCCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTGAGCGGAAGCCT
GCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCACGTGCACAACGC
CAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCTCCCGCCACCTCAT
GAAGGTCAAGCTCAAGAAGTACACCAAGTTCTCTTCGTGCGCGACCCCTTCGTGCGCCTGAT
CTCCGCCTTCCGCAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCGCAAGTTCGCCGTGCC
CATGCTGCGGCTGTACGCCAACCACACAGCCTGCCCGCCTCGGCGCGCGAGGCCTTCCGCGC
TGGCCTCAAGGTGTCCTTCGCCAAGTTTCATCCAGTACCTGCTGGACCCGCACACGGAGAAGCT
GGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCACCCGTGCCAGATCGACTA
CGACTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGCAGCTGCTGCAGCTACTCCA
GGTGGACCGGCAGCTCCGCTTCCCCCGAGCTACCGGAACAGGACCGCCAGCAGCTGGGAGGA
GGACTGGTTGCCAAGATCCCCCTGGCCTGGAGGCAGCAGCTGTATAAACTCTACGAGGCCGA
CTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCCGAGACTTGAAAGCTTTCGCGTTG
CTTTTTCTCGCGTGCCTGGAACCTGACGCACGCGCACTCCAGTTTTTTTATGACCTACGATTT
TGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCCATTGAGTACTGTATCGATATTGTT
TTTTAAGATTAATATATTTTCAGGTATTTAATACGA

418/550

FIGURE 418

MTKARLFRLWLVLGSVFMILLIIIVYWDSAGAAHFYLHTSFSPHTGPPLPTPGPDRDRELTAD
SDVDEFLDKFLSAGVKQSDLPRKETEQQPAPGSMEE SVRGYDWSPRDARRSPDQGRQQAERRS
VLRGFCANSSLAFTKERAFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKRMIVLSGSL
LHRGAPYRDPLRIPREHVHNASAHLTFNKFWRRYGKLSRHLMKVKLKKYTKFLFVRDPFVRLI
SAFRSKFELENEEFYRKFAVPMLRLYANHTSLPASAREAFRAGLKVSFANFIQYLLDPHTEKL
APFNEHWRQVYRLCHPCQIDYDFVGKLETLDDEAAQLLQLLQVDRQLRFPPSYRNRTASSWEE
DWFAKIPLAWRQQLYKLYEADFVLFGYPKPENLLRD

Important features:**Signal peptide:**

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

419/550

FIGURE 419

GGCACGAGGCTGAACCCAGCCGGCTCCATCTCAGCTTCTGGTTTCTAAGTCCATGTGCCAAAG
GCTGCCAGGAAGGAGACGCCTTCCTGAGTCCTGGATCTTTCTTCCTTCTGGAAATCTTTGACT
GTGGGTAGTTATTTATTTCTGAATAAGAGCGTCCACGCATCATGGACCTCGCGGGACTGCTGA
AGTCTCAGTTCCTGTGCCACCTGGTCTTCTGCTACGTCTTTATTGCCTCAGGGCTAATCATCA
ACACCATTACAGCTCTTCACTCTCCTCCTCTGGCCCATTAACAAGCAGCTCTTCCGGAAGATCA
ACTGCAGACTGTCCTATTGCATCTCAAGCCAGCTGGTGATGCTGCTGGAGTGGTGGTCGGGCA
CGGAATGCACCATCTTACGGACCCGCGCGCTACCTCAAGTATGGGAAGGAAAATGCCATCG
TGGTTCTCAACCACAAGTTTGAAATTGACTTTCTGTGTGGCTGGAGCCTGTCCGAACGCTTTG
GGCTGTTAGGGGGCTCCAAGGTCCTGGCCAAGAAAGAGCTGGCCTATGTCCCAATTATCGGCT
GGATGTGGTACTTCACCGAGATGGTCTTCTGTTTCGCGCAAGTGGGAGCAGGATCGCAAGACGG
TTGCCACCAGTTTGCAGCACCTCCGGGACTACCCCGAGAAGTATTTTTCTGATTCACTGTG
AGGGCACACGGTTCACGGAGAAGAAGCATGAGATCAGCATGCAGGTGGCCCGGGCCAAGGGG
TGCTCGCCTCAAGCATCACCTGTTGCCACGAACCAAGGGCTTCGCCATCACCGTGAGGAGCT
TGAGAAATGTAGTTTCAGCTGTATATGACTGTACACTCAATTCAGAAATAATGAAAATCCAA
CACTGCTGGGAGTCCTAAACGGAAAGAAATACCATGCAGATTTGTATGTTAGGAGGATCCCAC
TGGAAGACATCCCTGAAGACGATGACGAGTGCTCGGCCTGGCTGCACAAGCTCTACCAGGAGA
AGGATGCCTTTCAGGAGGAGTACTACAGGACGGGCACCTTCCCAGAGACGCCCATGGTGCCCC
CCCGGCGGCCCTGGACCCTCGTGAAGTGGCTGTTTTGGGCCTCGCTGGTGCTCTACCCTTTCT
TCCAGTTCCTGGTCAGCATGATCAGGAGCGGGTCTTCCCTGACGCTGGCCAGCTTCATCCTCG
TCTTCTTTGTGGCCTCCGTGGGAGTTCGATGGATGATTGGTGTGACGGAATTGACAAGGGCT
CTGCCTACGGCAACTCTGACAGCAAGCAGAACTGAATGACTGACTCAGGGAGGTGTCACCAT
CCGAAGGGAACCTTGGGGAAGTGGTGGCCTCTGCATATCCTCCTTAGTGGGACACGGTGACAA
AGGCTGGGTGAGCCCCCTGCTGGGCACGGCGGAAGTCACGACCTCTCCAGCCAGGGAGTCTGGT
CTCAAGGCCGGATGGGGAGGAAGATGTTTTGTAATCTTTTTTTCCCATGTGCTTTAGTGGGC
TTTGGTTTTCTTTTTGTGCGAGTGTGTGTGAGAATGGCTGTGTGGTGAGTGTGAAGTTGTTC
TGTGATCATAGAAAGGTATTTTAGGCTGCAGGGGAGGGCAGGGCTGGGGACCGAAGGGGACA
AGTTCCCCTTTCATCCTTTGGTGCTGAGTTTTCTGTAACCCTTGGTTGCCAGAGATAAAGTGA
AAAGTGCTTTAGGTGAGATGACTAAATTATGCCTCCAAGAAAAAAAATTAAGTGCTTTTCT
GGGTCAAAAAAAAAA

420/550

FIGURE 420

MDLAGLLKSQFLCHLVFCYVFIA SGLIINTIQLFTLLLWPINKQLFRKINCRLSYCISSQLVM
LLEWWSGTECTIFTDPRAYLKYGKENAIVVLNHKFEIDFLCGWSLSERFGLLGSKVLAKKEL
AYVPIIGWMWYFTMVFC SRKWEQDRKTVATSLQHLDYPEKYFFLIHCEGTRFTEKKHEISM
QVARAKGLPRLKHHLLPRTKGF AITVRSLRNVVSAVYDCTLNFRNNENPTLLGVLNGKKYHAD
LYVRRIPLEDIPEDDDECSAWLHKLYQEKDAFQEEYYRTGTFPETPMVPPRRPWTLVNWLFWA
SLVLYPFFQFLVSMIRSGSSLT LASFILVFFVASVGVRWMIGVTEIDKGSAYGNSDSKQKLND

421/550

FIGURE 421

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGGTGCCTGCATC
GCC**ATG**GACACCACCAGGTACAGCAAGTGGGGCGGCAGCTCCGAGGAGGTCCCCGAGGGCCC
TGGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTCTTCTTGGCCCTGGCTGTCCTGGTCACC
ACAGTCCTTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCGCGGCG
CTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCTCGAAGCAGACGGCGGCGCTGGGTGCC
CTGAAGGAGGAGGTTCGGAGACTGCCACAGCTGCTGCTCGGGGACGCAGGCGCAGCTGCAGACC
ACGCGCGCGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGCGCCCTGCGGGAA
CTGCGTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGGCAGGGGCCGTGAGGACGTCCGCACT
GAGCTGTTCCGGGCGCTGGAGGCCGTGAGGCTCCAGAACAACTCCTGCGAGCCGTGCCCCACG
TCGTGGCTGTCCTTCGAGGGCTCCTGCTACTTTTTCTCTGTGCCAAAGACGACGTGGGCGGCG
GCGCAGGATCACTGCGCAGATGCCAGCGCGCACCTGGTGATCGTTGGGGGCCTGGATGAGCAG
GGCTTCCTCACTCGGAACACGCGTGGCCGTGGTTACTGGCTGGGCCTGAGGGCTGTGCGCCAT
CTGGGCAAGGTTTCAGGGCTACCAGTGGGTGGACGGAGTCTCTCTCAGCTTCAGCCACTGGAAC
CAGGGAGAGCCCAATGACGCTTGGGGGCGCGAGAAGTGTGTCATGATGCTGCACACGGGGCTG
TGGAACGACGCACCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAGAAAAGGCACAACTGC
TGACCCCGCCAGTGCCCTGGAGCCGCGCCCATTTGCAGCATGTCGTATCCTGGGGGCTGCTCA
CCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTTTTCTTCCTCATCCACCGCTGCTGAG
TCTCAGAAACACTTGGCCCAACATAGCCCTGTCCAGCCCAGTGCCTGGGCTCTGGGACCTCCA
TGCCGACCTCATCCTAACTCCACTCACGCAGACCCAACCTAACCTCCACTAGCTCCAAAATCC
CTGCTCCTGCGTCCCCGTGATATGCCTCCACTTCTCTCCCTAACCAAGGTTAGGTGACTGAGG
ACTGGAGCTGTTTGGTTTTCTCGCATTTTCCACCAAACCTGGAAGCTGTTTTTGCAGCCTGAGG
AAGCATCAATAAATATTTGAGAAATGAAAAA

422/550

FIGURE 422

MDTTRYSKWGGSSSEVPGGPWGRVHWSRRPLFLALAVLVTTVLWAVILSILLSKASTERAAL
LDGHDLLRTNASKQTAALGALKEEVGDCHSCCSGTQAQLQTTRAEELGEAQAKLMEQESALREL
RERV TQGLAEAGRGREDVRTELFRAL EAVRLQNNSCEPCPTSWLSFEGSCYFFSVPKTTWAAA
QDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLSF SHWNQ
GEPNDAWGRENCVMMMLHTGLWNDAPCDSEKDGWICEKRHNC

Important features:**Type II transmembrane domain:**

amino acids 31-54

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287

423/550

FIGURE 423

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCCTTGAAAAACCCGGCGGGG
GAGCGAGGCTGCGGGCCGGCCGCTGCCCTTCCCCACACTCCCCGCCGAGAAGCCTCGCTCGGC
GCCCAAC**ATG**GCGGGTGGGCGCTGCGGCCCCGAGCTAACGGCGCTCCTGGCCGCTGGATCGC
GGCTGTGGCGGCGACGGCAGGCCCCGAGGAGGCCGCGCTGCCGCCGGAGCAGAGCCGGGTCCA
GCCCATGACCGCCTCCAACCTGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATTTTACGC
CCCATGGTGTCCATCCTGCCAGCAGACTGATTCAGAATGGGAGGCTTTTGCAAAGAATGGTGA
AATACTTCAGATCAGTGTGGGGAAGGTAGATGTCATTCAAGAACCAGGTTTGAGTGGCCGCTT
CTTTGTCACTACTCTCCAGCATTTTTTCATGCAAAGGATGGGATATTCCGCCGTTATCGTGG
CCAGGAATCTTCGAAGACCTGCAGAATTATATCTTAGAGAAGAAATGGCAATCAGTCGAGCC
TCTGACTGGCTGGAAATCCCCAGCTTCTCTAACGATGTCTGGAATGGCTGGTCTTTTTAGCAT
CTCTGGCAAGATATGGCATCTTCACAACATTTACAGTGACTCTTGGAATCCTGCTTGGTG
TTCTTATGTGTTTTTCGTCTAGCCACCTTGGTTTTTGGCCTTTTTATGGGTCTGGTCTTGGT
GGTAATATCAGAATGTTTCTATGTGCCACTTCCAAGGCATTTATCTGAGCGTTCTGAGCAGAA
TCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCGGAGGAGGAAAAAGATCTTG
TTCAAATGAAGAAGAAAACAAAGACAGCCTTGTAGATGATGAAGAAGAGAAAGAAGATCTTGG
CGATGAGGATGAAGCAGAGGAAGAAGAGGAGGAGGACAACCTGGCTGCTGGTGTGGATGAGGA
GAGAAGTGAGGCCAATGATCAGGGGCCCCCAGGAGAGGACGGTGTGACCCGGGAGGAAGTAGA
GCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAACCCTGCCAGCTGACACAGAGGTGGTGGGA
AGACTCCTTGAGGCAGCGTAAAAGTCAGCATGCTGACAAGGGACTG**TAG**ATTTAATGATGCGT
TTTCAAGAATACACACCAAAACAATATGTCAGCTTCCCTTTGGCCTGCAGTTTGTACCAAATC
CTTAATTTTTCTGAATGAGCAAGCTTCTCTTAAAGATGCTCTCTAGTCATTTGGTCTCATG
GCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGTGACAATCAGGATATAGAAAAAC
AAACGTAGTGTTGGGATCTGTTTGGAGACTGGGATGGGAACAAGTTCATTTACTTAGGGGTCA
GAGAGTCTCGACCAGAGGAGGCCATTCCAGTCCTAATCAGCACCTTCCAGAGACAAGGCTGC
AGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCTCCTGAGCATCCCCAAAGTGTAACGT
AGAAGCCTTGATCCTTTTTCTTGTTAAAGTATTTATTTTTGTCAAATTGCAGGAAACATCAG
GCACCACAGTGATGAAAAATCTTTCACAGCTAGAAATTGAAAGGGCCTTGGGTATAGAGAGC
AGCTCAGAAGTCATCCCAGCCCTCTGAATCTCCTGTGCTATGTTTTATTTCTTACCTTTAATT
TTTCCAGCATTTCCACCATGGGCATTTCAGGCTCTCCACACTCTTCACTATTATCTCTTGGTCA
GAGGACTCCAATAACAGCCAGGTTTACATGAACTGTGTTTGTTCATTCTGACCTAAGGGGTTT
AGATAATCAGTAACCATAACCCCTGAAGCTGTGACTGCCAAACATCTCAAATGAAATGTTGTG
GCCATCAGAGACTCAAAGGAAGTAAGGATTTTACAAGACAGATTAAAAAAAATTGTTTTGT
CCAAAATATAGTTGTTGTTGATTTTTTTTTTAAGTTTTCTAAGCAATATTTTTCAAGCCAGAAG
TCCTCTAAGTCTTGCCAGTACAAGGTAGTCTTGTGAAGAAAAGTTGAATACTGTTTTGTTTTT
ATCTCAAGGGGTTCCCTGGGTCTTGAACACTTTAATAATAACTAAAAAACCACTTCTGATTT
TCCTTCAGTGATGTGCTTTTGGTGAAAGAATTAATGAACCTCCAGTACCTGAAAGTGAAAGATT
TGATTTTGTTCATCTCTGTAATCTTCCAAGAATTATATCTTTGTAAATCTCTCAATACT
CAATCTACTGTAAGTACCCAGGGAGGCTAATTTCTTT

424/550

FIGURE 424

MAGGRCGPQLTALLAAWIAAVAATAGPEEAALPPEQSRVQPMTASNWTLVMEGEWMLKFYAPW
CPSCQQT DSEWEAF AKNGEILQISVGKVDVIQEPLSGRFFVTTLPAFFHAKDGI FRRYRGPG
IFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAWCSY
VFFVIATLVFGLFMGLVLVVISECFYVPLPRHLSEQRNRSEEAHRAEQLQDAEEEKDDSN
EEENKDSLVDDEEEKEDLGDEDEAEEDNLAAGVDEERSEANDQGPPGEDGVTREEVEPE
EAEEGISEQPCPADTEVVEDSLRQRKSHADKGL

Important features:**Signal peptide:**

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

425/550

FIGURE 425

GAGGAACCTACCGGTACCGGCCGCGCTGGTAGTCGCGGGTGTGGCTGCACCTACCAATCCCGTGCGCCGCGG
CTGGGCCGTCGGAGAGTGCGTGTGCTTCTCTCCTGCACGCGGTGCTTGGGCTCGGCCAGGCGGGGTCCGCCGCCA
GGGTTTGAGGATGGGGAGTAGCTACAGGAAGCGACCCGCGATGGCAAGGTATATTTTTGTGGAATGAAAAGGA
AGTATTAGAAATGAGCTGAAGACCATTACAGATTAATATTTTTGGGGACAGATTTGTGATGCTTGATTCACCT
TGAAGTAATGTAGACAGAAGTTCTCAAATTTGCATATTACATCAACTGGAACCAGCAGTGAATCTTAATGTTTAC
TTAAATCAGAACTTGCATAAGAAAGAGAATGGGAGTCTGGTTAAATAAAGATGACTATATCAGAGACTTGAAAAG
GATCATTCTCTGTTTTCTGATAGTGATATGGCCATTTTAGTGGGCACAGATCAGGATTTTTACAGTTTACTTGG
AGTGTCCAAAACGCAAGCAGTAGAGAAATAAGACAAGCTTTCAAGAAATTGGCATTGAAGTTACATCCTGATAA
AAACCCGAATAACCCAAATGCACATGGCGATTTTTTAAAAATAAATAGAGCATATGAAGTACTCAAAGATGAAGA
TCTACGGAAAAAGTATGACAAATATGGAGAAAAGGGACTTGAGGATAATCAAGGTGGCCAGTATGAAAGCTGGAA
CTATTATCGTTATGATTTTTGGTATTATGATGATGATCCTGAAATCATAACATTGGAAAGAGAGAATTTGATGC
TGCTGTTAATCTGGAGAACTGTGGTTGTAAATTTTTACTCCCCAGGCTGTTACACTGCCATGATTTAGCTCC
CACATGGAGAGACTTTGCTAAAGAAGTGGATGGTTACTTTCGAATTGGAGCTGTTAACTGTGGTGATGATAAAT
GCTTTGCCGAATGAAAGGAGTCAACAGCTATCCAGTCTCTCTTCATTTTTTCGGTCTGGAATGGCCCCAGTGAAATA
TCATGGAGACAGATCAAAGGAGAGTTTAGTGAGTTTTGCAATGCAGCATGTTAGAAGTACAGTGACAGAACTTTG
GACAGGAAATTTTGTCAACTCCATCAAACGCTTTTGGCTGCTGGTATTGGCTGGCTGATCACTTTTTGTTCAAA
AGGAGGAGATTGTTTGACTTCACAGACACGACTCAGGCTTAGTGGCATGTTGTTTTCTCAACTCATTGGATGCTAA
AGAAATATATTTGGAAGTAATACATAATCTTCCAGATTTTGAACACTTTTCGGCAAACACACTAGAGGATCGTTT
GGCTCATCATCGGTGGCTGTTATTTTTTCATTTTTGGAAAAAATGAAAATTCAAATGATCCTGAGCTGAAAAA
AAAACTCTACTTAAAAATGATCATATTCAAGTTGGCAGGTTTGAAGTTTCTCTGCACCAGACATCTGTAGTAA
TCTGTATGTTTTTTCAGCCGCTCTCTAGCAGTATTTAAAGGACAAGGAACCAAAGAATATGAAATTCATCATGGA
GAAGATTCTATATGATATACTTGCCCTTTGCCAAAGAAAGTGTGAATCTCATGTTACCACGCTTGGACCTCAAA
TTTTCTGCCAATGACAAAGAACCATGGCTTGTGTTCTTTGCCCCCTGGTGCCACCATGTGAGCTTACT
ACCGAGTTACGAAGAGCATCAAATCTTTATGTTGATCTTATGTTAGCTTAAGTTTGGTACACTAGATTGTACAGTTCATGA
GGGACTCTGTAACATGTATAACATTCAAGGCTTATCCAACAACAGTGGTATTCAACCAGTCCAACATTTCATGAGTA
TGAAGGACATCACTCTGCTGAACAAATCTTGGAGTTCATAGAGGATCTTATGAATCCTTCAGTGGTCTCCCTTAC
ACCCACCACCTTCAACGAAGTGTACACAAAGAAAACACAACGAAGTCTGGATGGTTGATTTCTATTCTCCGTG
GTGTCATCCTTGCCAAAGTCTTAATGCCAGAATGGAAAGTGGCCCGGACATTAACTGGACTGATCAACGTGGG
CAGTATAGATTGCCAAAGCATCATTTCTTTTTGTGCCAGGAAAACGTTCAAAGATACCCCTGAGATAAGATTTTT
TCCCCCAAATCAAATAAAGCTTATCAGTATCACAGTTACAATGGTTGGAATAGGGATGCTTATCCCTGAGAAT
CTGGGGTCTAGGATTTTTACCTCAAGTATCCACAGATCTAACACCTCAGACTTTTCAGTGAAGAAAGTTCTACAAGG
GAAAACTCATTGGGTGATTGATTTCTATGCTCCTTGGTGTGGACCTTGCCAGAATTTTGTCCAGAATTTGAGCT
CTTGGCTAGGATGATTAAGGAAAAGTGAAGCTGGAAAAGTAGACTGTCAAGCTTATGCTCAGACATGCCAGAA
AGCTGGGATCAGGGCTTATCCAAGTGTAAAGTTTATTTCTACGAAAGAGCAAAGAGAAATTTTCAAGAAGAGCA
GATAAATACCAGAGATGCAAAAGCAATCGCTGCCTTAATAAGTGAAGAAATTTGGAACTCTCCGAAATCAAGGCCAA
GAGGAATAAGGATGAACCTTGGATAATGTTGAAGATGAAGAAAAAGTTTAAAGAAATTTCTGACAGATGACATCAG
AAGACACCTATTTAGAATGTTACATTTATGATGGGAATGAATGAACATTATCTTAGACTTGCAAGTTGTACTGCCA
GAATTATCTACAGCACTGGTGTAAAGAAGGGTCTGCAAACTTTTTCTGTAAGGGCCGGTTTATAAATATTTTA
GACTTTGCAAGCTATAATATATGTTTACACATGAGAACAAGAATAGAGTCATCATGTATTCTTTGTTATTTGCT
TTTAACAACCTTTAAAAAATATTAACGATTCTTAGCTCAGAGCCATACAAAAGTAGGCTGGATTTCAGTCCATG
GACCATAGATTGCTGTCCCCCTCGACGGACTTATAATGTTTCAGGTGGCTGGCTTGAACATGAGTCTGCTGTGCT
ATCTACATAAATGTCTAAGTTGTATAAAGTCCACTTTCCCTTCACGTTTTTGGCTGACCTGAAAAGAGGTAAC
TAGTTTTTGGTCACTTGTTCTCTTAAAAATGCTATCCCTAACCATATATTTATATTTTCGTTTTAAAAACACCCAT
GATGTGGCACAGTAAACAAACCCCTGTTATGCTGTATTATTATGAGGAGATTCTTCATTGTTTTCTTCTCTCA
AAGGTTGAAAAAATGCTTTTTAATTTTTTACAGCCGAGAAACAGTGCAGCAGTATATGTGCACACAGTAAGTACAC
AAATTTGAGCAACAGTAAGTGCACAAATCTGTAGTTTGTGTATCATCCAGGAAAACCTGAGGGAAAAAATTA
TAGCAATTAAGTGGGATTTGTAGAGTATCCTAAATATGTTATCAAGTATTTAGAGTTCTATATTTTAAAGATATA
TGTGTTCTATGATTTTCTGAAATGCTTTTATAGAAATTTTCCCACTGATAGTTGATTTTGGAGGCATCTAATAT
TTACATATTTGCTTCTGAACCTTTGTTTTGACCTGTATCCTTTATTTACATTGGGTTTTTCTTTTCATAGTTTTGG
TTTTTCACTCCTGTCCAGTCTATTTATTATTCAAATAGGAAAAATTACTTTACAGGTTGTTTTACTGTAGCTTAT
AATGATACTGTAGTTATTCCAGTTACTAGTTTACTGTGAGAGGGCTGCCTTTTTTCAGATAAATATTGACATAATA
ACTGAAGTTATTTTTATAAGAAAATCAAGTATATAAATCTAGGAAAGGGATCTTCTAGTTTCTGTGTTGTTTGA
CTCAAAGAATCACAATTTGTCAGTAACATGTAGTTGTTTATGTTTATAAATTCAGAGTGTACAGAATGGTAAAAAT
CCAATCAGTCAAAGAGGTCATGAATTAAGGCTTGCAACTTTTTCAAAAAA

426/550

FIGURE 426

MGVWLNKDDYIRDLKRIILCFLIVYMAILVGTDQDFYSLLGVSKTASSREIRQAFKKLALKLH
PDKNPNNPNAHGDFLKNRAYEVLKDEDLRKKYDKYGEKGLEDNQGGQYESWNNYYRYDFGIYD
DDPEIITLERREFDAAVNSGELWVFNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNCGDD
RMLCRMKGVNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNSIQTA
FAAGIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSANTLEDR
LAHHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICSPLYVFQPSLAVFK
GQGTKEYEIHGKKILYDILAFAKESVNSHVTTLGPQNFPANDKEPWLVDFFAPWCPPCRALL
PELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIEYEGHHSAEQILEFI
EDLMNPSVVSLTPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMARTLTGLINVGS
IDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLGFLPQVSTDLT
PQTFSEKVLQGNHWWIDFYAPWCGPCQNFAPFELLARMIKGKVKAGKVDCQAYAQTCQKAG
IRAYPTVKFYFYERAKRNFQEEQINTRDAKAI AALISEKLETLRNQGRNKDEL

Important features:**Endoplasmic reticulum targeting sequence.**

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

427/550

FIGURE 427

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGCA
CGGTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTTTCTCCCTCTTGA
GTCCTTCTGAGATGATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTGCGGATGGTAG
CGGCGGCTCTCGGCGGCCACCCTCTGCTGGGAGTGAGCGCCACCTTGAACTCGGTTCTCAATT
CCAACGCTATCAAGAACCTGCCCCACCGCTGGGCGGCGCTGCGGGGCACCCAGGCTCTGCAG
TCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACAACCTACC
AGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGTCCCACCC
GCGGAGGGGACGCAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACGCTGCATGC
GTCACGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTTCTGATCAAA
ATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGATCATAGCACCT
TGGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAAAGGACAAGAAG
GTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTAGACACTTCTGGT
CCAAGATCTGTAAACCTGTCCTGAAAGAAGGTCAAGTGTGTACCAAGCATAGGAGAAAAGGCT
CTCATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTCTTGCCGGATACAGA
AAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTGAGAGACACTTAAACCAGCT
ATCCAAATGCAGTGAACCTCTTTTATATAATAGATGCTATGAAAACCTTTTATGACCTTCATC
AACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCATTCCAATAACACCTTCCA
AAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATGGAACCTCCCCTGTGATTGCAGTAAATTACT
GTATTGTAAATTCTCAGTGTGGCACTTACCTGTAAATGCAATGAACTTTTAATTATTTTCT
AAAGGTGCTGCACTGCCTATTTTCTCTTGTATGTAAATTTTGTACACATTGATTGTTAT
CTTGACTGACAAATATTCTATATTGAACTGAAGTAAATCATTTTCAGCTTATAGTTCTTAAAAG
CATAACCTTTACCCATTTAATTCTAGAGTCTAGAACGCAAGGATCTCTTGAATGACAAAT
GATAGGTACCTAAAATGTAACATGAAAATACTAGCTTATTTTCTGAAATGTACTATCTTAATG
CTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTTGAAATAAAATTTAACATTTAAAAAA
AAAAAA

428/550

FIGURE 428

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSAA
PGILYPGGNKYQTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRKRRCMRHAM
CCPGNYCKNGICVSSDQNHFRGEIETITESFGNDHSTLDGYSRRTLSSKMYHTKGQEGSVC
LRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQKDDH
QASNSSRLHTCQRH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

429/550

FIGURE 429

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCCGCTCCCGGAGCCCAGCCC
TTTCCTAACCCAACCAACCTAGCCCAGTCCCAGCCGCCAGCGCCTGTCCCTGTCACGGACCC
CAGCGTTACCAATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCT
CCTGGTAACTTGGGTTTTTACTCCTGTAACAACTGAAATAACAAGTCTTGCTACAGAGAATAT
AGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTTATGCTGACTGGTGTCTGTTT
CAGTCAGATGTTGCATCCAATTTTTGAGGAAGCTTCCGATGTCATTAAGGAAGAATTTCCAAA
TGAAAATCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATA
CAGGATAAGCAAATACCCAACCCCTCAAATTGTTTCGTAATGGGATGATGATGAAGAGAGAATA
CAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGCAACAAAAAAGTGACCCCAT
TCAAGAAATTCGGGACTTAGCAGAAATCACCCTCTTGATCGCAGCAAAGAAATATCATTGG
ATATTTTGAGCAAAGGACTCGGACAACTATAGAGTTTTTGAACGAGTAGCGAATATTTTGCA
TGATGACTGTGCCTTTCTTTCTGCATTTGGGGATGTTTCAAACCGGAAAGATATAGTGGCGA
CAACATAATCTACAAACCACCAGGGCATTCTGCTCCGGATATGGTGTACTTGGGAGCTATGAC
AAATTTTGATGTGACTTACAATTGGATTCAAGATAAATGTGTTCCTCTTGTCGAGAAATAAC
ATTTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCATACTCTTTCACATGAA
AGAAGATACAGAAAGTTTAGAAATATTCCAGAATGAAGTAGCTCGGCAATTAATAAGTGAAAA
AGGTACAATAAACTTTTTACATGCCGATTGTGACAAATTTAGACATCCTCTTCTGCACATACA
GAAACTCCAGCAGATTGTCCTGTAATCGCTATTGACAGCTTTAGGCATATGTATGTGTTTGG
AGACTTCAAAGATGTATTAATTCCTGGAAAACCTCAAGCAATTCGTATTTGACTTACATTCTGG
AAAACCTGCACAGAGAATTCCATCATGGACCTGACCCAACTGATACAGCCCCAGGAGAGCAAGC
CCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAACTAGCACCCAGTGAATATAG
GTATACTCTATTGAGGGATCGAGATGAGCTTTAAAAAACTTGAAAAACAGTTTGTAAGCCTTTC
AACAGCAGCATCAACCTACGTGGTGGAAATAGTAAACCTATATTTTCATAATTCTATGTGTAT
TTTTATTTTGAATAAACAGAAAGAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAA

430/550

FIGURE 430

MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQM
LHPIFEEASDVIKEEFPNENQVVFARVDCDQHS DIAQRYRISKYPTLKLFRNGMMMKREYRGQ
RSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKSDNYRVFERVANILHDDC
AFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNW IQDKCVPLVREITFEN
GEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLHIQKTP
ADCPVIAIDSFRHMYVFGDFKDVLIPGKLGKQFVFDLHSGKLHREFHHGPDPTDTAPGEQAQDV
ASSPPESSFQKLAPSEYRYTLLRDRDEL

Important features:

Signal peptide:

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

431/550

FIGURE 431

GAGCAGGACGGAGCCATGGACCCCGCCAGGAAAGCAGGTGCCCAGGCCATGATCTGGACTGCA
GGCTGGCTGCTGCTGCTGCTGCTTCGCGGAGGAGCGCAGGCCCTGGAGTGCTACAGCTGCGTG
CAGAAAGCAGATGACGGATGCTCCCCGAACAAGATGAAGACAGTGAAGTGCGCGCCGGGCGTG
GACGTCTGCACCGAGGCCGTGGGGGCGGTGGAGACCATCCACGGACAATTCTCGCTGGCAGTG
CGGGGTTGCGGTTCTGGGACTCCCCGGCAAGAATGACCGCGGCCTGGATCTTCACGGGCTTCTG
GCGTTCATCCAGCTGCAGCAATGCGCTCAGGATCGCTGCAACGCCAAGCTCAACCTCACCTCG
CGGGCGCTCGACCCGGCAGGTAATGAGAGTGCATACCCGCCCAACGGCGTGGAGTGCTACAGC
TGTGTGGGCCCTGAGCCGGGAGGCGTGCCAGGGTACATCGCCGCCGGTTCGTGAGCTGCTACAAC
GCCAGCGATCATGTCTACAAGGGCTGCTTCGACGGCAACGTCACCTTGACGGCAGCTAATGTG
ACTGTGTCCTTGCCTGTCCGGGGCTGTGTCCAGGATGAATTCTGCACTCGGGATGGAGTAACA
GGCCCAGGGTTCACGCTCAGTGGCTCCTGTTGCCAGGGGTCCCGCTGTAACCTCTGACCTCCGC
AACAAGACCTACTTCTCCCCCTCGAATCCCAACCCTTGTCCGGCTGCCCCCTCCAGAGCCCACG
ACTGTGGCCTCAACCACATCTGTCACCACTTCTACCTCGGCCCCAGTGAGACCCACATCCACC
ACCAAACCCATGCCAGCGCCAACCAGTCAGACTCCGAGACAGGGAGTAGAACACGAGGCCTCC
CGGGATGAGGAGCCAGGTTGACTGGAGGCGCCGCTGGCCACCAGGACCGCAGCAATTCAGGG
CAGTATCCTGCAAAAGGGGGGGCCCCAGCAGCCCCATAATAAAGGCTGTGTGGCTCCCACAGCT
GGATTGGCAGCCCTTCTGTTGGCCGTGGCTGCTGGTGTCTTACTGTGAGCTTCTCCACCTGGA
AATTTCCCTCTCACCTACTTCTCTGGCCCTGGGTACCCCTCTTCTCATCACTTCCTGTTCCCA
CCACTGGACTGGGCTGGCCAGCCCCCTGTTTTTCCAACATTCCCCAGTATCCCCAGCTTCTGC
TGCGCTGGTTTGCGGCTTTGGGAAATAAAATACCGTTGTATATATTCTGCCAGGGGTGTTCTA
GCTTTTTGAGGACAGCTCCTGTATCCTTCTCATCCTTGTCTCTCCGCTTGTCCTCTTGTGATG
TTAGGACAGAGTGAGAGAAGTCAGCTGTACGGGGAAGGTGAGAGAGAGGATGCTAAGCTTCC
TACTCACTTTCTCCTAGCCAGCCTGGACTTTGGAGCGTGGGGTGGGTGGGACAATGGCTCCCC
ACTCTAAGCACTGCCTCCCCTACTCCCCGCATCTTTGGGGAATCGGTTCCCCATATGTCTTCC
TTACTAGACTGTGAGCTCCTCGAGGGGGGGCCCCGGTACCCAATTCGCCCTATAGTGAGTCGTA

432/550

FIGURE 432

MDPARKAGAQAMIWTAGWLLLLLLRGGAQALECYSCVQKADDGCSPNKMKTVKCAPGVDVCTE
AVGAVETIHGQFSLAVRGCGSGLPGKNDRGLDLHGLLAFIQLQQAQDRCAKLNLTSRALDP
AGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGCDFGNVTLTAANVTVSLP
VRGCVQDEFCTRDGVTGPGFTLSGSCCQGSRCNSDLRNKTYFSPRIPLVRLPPPEPTTVAST
TSVTTSTSAPVRPTSTTKPMPAPTSQTPRQGVHEASRDEEPRLTGGAAGHQDRSNSGQYPAK
GGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

433/550

FIGURE 433

CGGGACTCGGGTGCTCCTCTGGGAGTCTCGGAGGGGACCGGCTGTGCAGACGCC**ATGGAGTT**
GGTGCTGGTCTTCTCTGCAGCCTGCTGGCCCCCATGGTCTGGCCAGTGCAGCTGAAAAGGA
GAAGGAAATGGACCCTTTTCATTATGATTACCAGACCCTGAGGATTGGGGGACTGGTGTTCGC
TGTGGTCTCTTCTCGGTTGGGATCCTCCTTATCCTAAGTCGCAGGTGCAAGTGCAGTTTCAA
TCAGAAGCCCCGGGCCCCAGGAGATGAGGAAGCCAGGTGGAGAACCTCATCACCGCCAATGC
AACAGAGCCCCAGAAGCAGAGAACTGAAGTGCAGCCATCAGGTGGAAGCCTCTGGAACCTGAG
GCGGCTGCTTGAACCTTTGGATGCAAAATGTCGATGCT**TAA**GAAAACCGGCCACTTCAGCAACA
GCCCTTTCCCCAGGAGAAGCCAAGAACTTGTGTGTCCCCACCCTATCCCCTCTAACACCATT
CCTCCACCTGATGATGCAACTAACACTTGCCTCCCCACTGCAGCCTGCGGTCTGCCCACCTC
CCGTGATGTGTGTGTGTGTGTGTGTGTGACTGTGTGTGTTTGCTAACTGTGGTCTTTGTGG
CTACTTGTGTGTGGATGGTATTGTGTTGTAGTGAACGTGGACTCGCTTTCCAGGCAGGG
GCTGAGCCACATGGCCATCTGCTCCTCCCTGCCCCCGTGGCCCTCCATCACCTTCTGCTCCTA
GGAGGCTGCTTGTGTTGCCCCGAGACCAGCCCCCTCCCCTGATTTAGGGATGCGTAGGGTAAGAGC
ACGGGCAGTGGTCTTTCAGTCGTCTTGGGACCTGGGAAGGTTTGCAGCACTTTGTTCATCATTCT
TCATGGACTCCTTTCACTCCTTTAAACAAAAACCTTGCTTCCTTATCCCACCTGATCCCAGTCT
GAAGGTCTCTTAGCAACTGGAGATACAAAGCAAGGAGCTGGTGAGCCCAGCGTTGACGTCAGG
CAGGCTATGCCCTTCCGTGGTTAATTTCTTCCAGGGGCTTCCACGAGGAGTCCCCATCTGCC
CCGCCCCCTTACAGAGCGCCCGGGGATTCCAGGCCCAGGGCTTCTACTCTGCCCCCTGGGGAAT
GTGTCCCCCTGCATATCTTCTCAGCAATAACTCCATGGGCTCTGGGACCCTACCCCTTCCAACC
TTCCCTGCTTCTGAGACTTCAATCTACAGCCCAGCTCATCCAGATGCAGACTACAGTCCCTGC
AATTGGGTCTCTGGCAGGCAATAGTTGAAGGACTCCTGTTCCGTTGGGGCCAGCACACCGGGA
TGGATGGAGGGAGAGCAGAGGCCTTTGCTTCTCTGCCTACGTCCCCTTAGATGGGCAGCAGAG
GCAACTCCCGCATCCTTTGCTCTGCCTGTGCGGTGGTCAGAGCGGTGAGCGAGGTGGGTTGGAG
ACTCAGCAGGCTCCGTGCAGCCCTTGGGAACAGTGAGAGGTTGAAGGTCATAACGAGAGTGGG
AACTCAACCCAGATCCCGCCCCCTCCTGTCCTCTGTGTTCCCGCGGAAACCAACCAAACCGTGC
GCTGTGACCCATTGCTGTTCTCTGTATCGTGATCTATCCTCAACAACAACAGAAAAAAGGAAT
AAAATATCCTTTGTTTCCT

434/550

FIGURE 434

MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRCKC
SFNQKPRAPGDDEEAQVENLITANATEPQKQRTQEVQPSGGSLWNLRRLLLEPLDANVDA

435/550

FIGURE 435

GGTCCTTA**ATG**GCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTGC
TGTCCGGCTGGTCCCGGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCACCGTCA
TCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGACTT
TTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAACTAA
ATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACTTACAG
AGCAACTGCGTGACATTCAGCTGGAGAATTACACACCCAAGGAACCCCTCACCCCTGCAGGCCAA
GGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGTTTCGATG
GGCAGATCTTCTCCTCTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCCTGGAGCCA
GAAAGATGAAAGAAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCCTTCCATTACTTCTCAA
TGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACCCCTGGAGCCAA
GTGCAGGAGCACCACTCGCCATGTCCTCAGGCACAACCCAACTCAGGGCCACAGCCACCACCC
TCATCCTTTGCTGCCTCCTCATCATCCTCCCCTGCTTCATCCTCCCTGGCATCTGAGGAGAGT
CCTTTAGAGTGACAGGTTAAAGCTGATACCAAAGGCTCCTGTGAGCACGGTCTTGATCAAAC
TCGCCCTTCTGTCTGGCCAGCTGCCACGACCTACGGTGTATGTCCAGTGGCCTCCAGCAGAT
CATGATGACATCATGGACCCAATAGCTCATTCACTGCCTTGATTCTTTTGCCAACAATTTTA
CCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTGGTGCTACCTGATGGAATTCCTGCA
CTTAAAGTTCTGGCTGACTAAACAAGATATATCATTTTCTTTCTTCTTTTTGTTTGGA^{AAA}
TCAAGTACTTCTTTGAATGATGATCTCTTTCTTGCAAATGATATTGTCAGTAAAATAATCACG
TTAGACTTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAGAATTTTAAATTATTTAAT
AAGAAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTATTGTTCTGTACTGATATTTAA
ATAAAGAGTTCTATTTCCCAAAAAAAAAAAAAAAAAA

436/550

FIGURE 436

MAAAAATKILLCLPLLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTFLH
YDCGNKTVTPVSPLGKKLNVTTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQARMS
CEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFSMGD
CIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

437/550

FIGURE 437

GTTCTCCTTTCCGAGCCAAAATCCCAGGCGATGGTGAATTATGAACGTGCCACACCATGAAGCTCTTGTGGCAGG
TAACTGTGCACCACCACACCTGGAATGCCATCCTGCTCCCGTTTCGTCTACCTCACGGCGCAAGTGTGGATTCTGT
GTGCAGCCATCGCTGCTGCCGCTCAGCCGGGCCCCAGAACTGCCCCCTCCGTTTGCTCGTGCAGTAACCAGTTCA
GCAAGGTGGTGTGCACGCGCCGGGGCCTCTCCGAGGTCCCGCAGGGTATTCCCTCGAACACCCGGTACCTCAACC
TCATGGAGAACAACATCCAGATGATCCAGGCCGACACCTTCCGCCACCTCCACCACCTGGAGGTCTGCAGTTGG
GCAGGAACCTCATCCGGCAGATTGAGGTGGGGGCCCTTCAACGGCCTGGCCAGCCTCAACACCCTGGAGCTGTTCCG
ACAACCTGGCTGACAGTCATCCCTAGCGGGGCCTTTGAATACCTGTCCAAGCTGCGGGAGCTCTGGCTTCGCAACA
ACCCCATCGAAAGCATCCCCCTTACGCCTTCAACCGGGTGCCCTCCCTCATGCGCCTGGACTTGGGGGAGCTCA
AGAAGCTGGAGTATATCTCTGAGGGAGCTTTTGAGGGGCTGTTCAACCTCAAGTATCTGAACCTTGGGCATGTGCA
ACATTAAAGACATGCCAATCTCACCCCCCTGGTGGGGCTGGAGGAGCTGGAGATGTCAGGGAACCACTTCCCTG
AGATCAGGCCTGGCTCCTTCCATGGCCTGAGCTCCCTCAAGAAGCTCTGGGTCTGAACTCACAGGTGAGCCTGA
TTGAGCGGAATGCTTTTGACGGGCTGGCTTCACTTGTGGAACCTCAACTTGGCCACAATAACCTCTCTTCTTTGC
CCCATGACCTCTTTACCCCGTGAGGTACCTGGTGGAGTTGCATCTACACCACAACCCCTTGGAACTGTGATTGTG
ACATTCTGTGGCTAGCCTGGTGGCTTCGAGAGTATATACCCACCAATTCCACCTGCTGTGGCCGCTGTCTATGCTC
CCATGCACATGCGAGGCCGTACCTCGTGGAGGTGGACCAGGCCTCCTTCCAGTGCTCTGCCCCCTTCATCATGG
ACGCACCTCGAGACCTCAACATTTCTGAGGGTCGGATGGCAGAACTTAAGTGTGCGACTCCCCCTATGTCTCCG
TGAAGTGGTTGCTGCCCAATGGGACAGTGCTCAGCCACGCCTCCCGCCACCCAAGGATCTCTGTCTCAACGACG
GCACCTTGAACCTTTTCCCACGTGCTGCTTTCAGACACTGGGGTGACACATGCATGGTGACCAATGTTGCAGGCA
ACTCCAACGCCTCGGCCTACCTCAATGTGAGCACGGCTGAGCTTAACACCTCCAACCTACAGCTTCTTCACCACAG
TAACAGTGGAGACCACGGAGATCTCGCCTGAGGACACAACGCGAAAGTACAAGCCTGTTCTTACCACGTCCACTG
GTTACCAGCCGGCATATACCACCTTACCACGGTGCTCATTGAGACTACCCGTGTGCCCAAGCAGGTGGCAGTAC
CCGCGACAGACACCACTGACAAGATGCAGACCAGCCTGGATGAAGTCATGAAGACCACCAAGATCATCATTTGGCT
GCTTTGTGGCAGTGACTCTGCTAGCTGCCGCCATGTTGATTGTCTTCTATAAACTTCGTAAGCGGCACCAGCAGC
GGAGTACAGTCACAGCCGCCCCGACTGTTGAGATAATCCAGGTGGACGAAGACATCCCAGCAGCAACATCCGCAG
CAGCAACAGCAGCTCCGTCCGGTGTATCAGGTGAGGGGGCAGTAGTGCTGCCCAACAATTCATGACCATATTAAC
ACAACACCTACAAACCAGCACATGGGGCCCACTGGACAGAAAACAGCCTGGGGAACCTCTCTGCACCCCACAGTCA
CCACTATCTCTGAACCTTATATAATTGAGACCATACCAAGGACAAGGTACAGGAACTCAAATATTGACTCCCCCT
CCCCCAAAAACTTATAAAATGCAATAGAATGCACACAAAGACAGCAACTTTTGTACAGAGTGGGGAGAGACTTT
TTCTTGTATATGCTTATATATTAAGTCTATGGGCTGGTTAAAAAAAACAGATTATATTAATAATTTAAAGACAAAA
AGTCAAAACA

438/550

FIGURE 438

MKLLWQVTVHHHTWNAILLPFVYLTAQVWILCAAIAAAASAGPQNCPSVCSCSNQFSKVVCTR
RGLSEVPQGIPSNTRYLNLMEENIQMIQADTFRHLHHLEVLQLGRNSIRQIEVGAFNGLASLN
TLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIPSYAFNRVPSLMRLDLGELKKLEYISE
GAFEGLFNLKYLNLGMCNIKDMPNLTPLVGLLEELEMSGNHFPEIRPGSFHGLSSLKKLWVMNS
QVSLIERNAFDGLASLVELNLAHNNLSSLPHDLFTPLRYLVELHLHNPWNCDICDILWLAWWL
REYIPTNSTCCGRCHAPMHMRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAELKCRTPP
MSSVKWLLPNGTVLSHASRHPRISVLNDGTLNFSHVLLSDTGVTTCMVTNVAGNSNASAYLNV
STAEINTSNYSFFTTVTVETTEISPEDTTRKYKPVPTTSTGYQPAYTTSTTVLIQTTRVPKQV
AVPATDTTDKMQTSLDEVMKTTKIIIGCFVAVTLLAAAMLIVFYKLRKRHQQRSTVTAARTVE
IIQVDEDI PAATSAAATAAPSGVSGEGAVVLPTIHDHINYNTYKPAHGAHW TENS LGNSLHPT
VTTISEPYIIQTHTKDKVQETQI

439/550

FIGURE 439

GTCGAATCCAAATCACTCATTGTGAAAGCTGAGCTCACAGCCGAATAAGCCACCATGAGGCTG
TCAGTGTGTCTCCTGATGGTCTCGCTGGCCCTTTGCTGCTACCAGGCCCATGCTCTTGCTCTGC
CCAGCTGTTGCTTCTGAGATCACAGTCTTCTTATTCTTAAGTGACGCTGCGGTAAACCTCCAA
GTTGCCAAACTTAATCCACCTCCAGAAGCTCTTGACGCCAAGTTGGAAGTGAAGCACTGCACC
GATCAGATATCTTTTAAGAAACGACTCTCATTGAAAAAGTCCTGGTGGAAATTAGTGAAAAAAT
GTGGTGTGTGACATGTAAAAATGCTCAACCTGGTTTCCAAAGTCTTTCAACGACACCCTGATC
TTCATAAAAAATTGTAAAGGTTTCAACACGTTGCTTTAATAAATCACTTGCCCTGC

440/550

FIGURE 440

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVNLQVAKLNPPPEALAAKLEVK
HCTDQISFKKRLSLKKSWWK

441/550

FIGURE 441

GAACATTTTTAGTTCCCAAGGAATGTACATCAGCCCCACGGAAGCTAGGCCACCTCTGGGATG
GGGTTGCTGGTTTAAACAAACGCCAGTCATCCTATATAAGGACCTGACAGCCACCAGGCACC
ACCTCCGCCAGGAAGTGCAGGCCACCTGTCTGCAACCCAGCTGAGGCCATGCCCTCCCCAGG
GACCGTCTGCAGCCTCCTGCTCCTCGGCATGCTCTGGCTGGACTTGGCCATGGCAGGCTCCAG
CTTCCTGAGCCCTGAACACCAGAGAGTCCAGCAGAGAAAGGAGTCGAAGAAGCCACCAGCCAA
GCTGCAGCCCCGAGCTCTAGCAGGCTGGCTCCGCCCGGAAGATGGAGGTCAAGCAGAAGGGGC
AGAGGATGAACTGGAAGTCCGGTTCAACGCCCCCTTTGATGTTGGAATCAAGCTGTCAGGGGT
TCAGTACCAGCAGCACAGCCAGGCCCTGGGGAAGTTTCTTCAGGACATCCTCTGGGAAGAGGC
CAAAGAGGCCCCAGCCGACAAGTGATCGCCCACAAGCCTTACTCACCTCTCTCTAAGTTTAGA
AGCGCTCATCTGGCTTTTCGCTTGCTTCTGCAGCAACTCCCACGACTGTTGTACAAGCTCAGG
AGGCGAATAAATGTTCAAACCTGTA

442/550

FIGURE 442

MPSPGTVCSLLLLGLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDGG
QAEGAEDELEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO

443/550

FIGURE 443

CGGCCACAGCTGGCATGCTCTGCCTGATCGCCATCCTGCTGTATGTCCTCGTCCAGTACCTCG
TGAACCCCGGGGTGCTCCGCACGGACCCAGATGTCAAGAATATGAACACGTGGCTGCTGTTTC
CTCCCCCTGTTCCCGGTGCAGGTGCAGACCCTGATAGTCGTGATCATCGGGATGCTCGTGCTC
CTGCTGGACTTTCTTGGCTTGGTGCACCTGGGCCAGCTGCTCATCTTCCACATCTACCTGAGT
ATGTCCCCCACCCTAAGCCCCGATCCCCCAAGGCTGGGTGGTCAGAGCTGCTCATCTTACA
CCTCTACTTGAGTATGTCCCTAACCCTGAGCCCCCACGCTGGGGCCAGAGTCTTTGTCCCC
CGTGTGCGCATGTGTTCAAGGTTCAGCCTCTCCAGAAGTGAGATCATGGACAAAAAGGGCAA
TCACAGGAAGAAATTAAATCCATGAGGACCCAGCAGGCCAGCAAGAAGCTGAACTCACGCCG
AGACCTGCAGGAGTGGTGCCAGGTGCTTGAAGTAACAAGTTTAAAATGTTTCAGAGACAATGGA
ATGGAATCTATTAGGCAAGAACAGGACATTATGAAATAAGGACAGGTGGACTTCCAAAAACAC
AAGTAGAAATTCTAACAATGAAATATATTACAGGCAGGTCACCCACTAACCAAACTGAAG
CGAGAGCTGTGGTCTTGCTTGGTCTCACAGTGGGCACAGCGGTAGGCGGTGAGTCATGTTGCT
GAACGACGGAGGGTAAACTCCCCAGCCCCAAGAAAACCTGTGTTGGAAGTAACAACAACCTCC
CTGCTCCTGGCACCAGCCGTTTTGGTCATGGTGGGCCAGCTGCAAAGCGTCTTCCATTCTCTG
GGCAGTGGTGGCCCCGAGGCTGTGGCCTCTCAGGGGGTTTTCTGTGGACACGGGCAGCAGAGTG
TGTCCAGGCCAGCCCCAAGAATGCCCTGCTCCTGACAGCTTGGCCAACCCCTGGTCAGGGCA
GAGGGAGTTGGGTGGGTGAGGCTCTGGGCTCACCTCCATCTCCAGAGCATCCCCTGCCTGCAG
TTGTGGCAAGAACGCCCAGCTCAGAATGAACACACCCACCAAGAGCCTCCTTGTTTCATAACC
ACAGGTTACCCTACAAACCACTGTCCCCACACAACCCTGGGGATGTTTTAAACACACACCTC
TAACGCATATCTTACAGTCACTGTTGTCTTGCTGAGGGTTGAATTTTTTTTAATGAAAGTGC
AATGAAAATCACTGGATTAAATCCTACGGACACAGAGCTGAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAA

444/550

FIGURE 444

MNTWLLFLPLFPVQVQTLIVVIIGMLVLLLDLGLVHLGQLLIFHIYLSMSPTLSPRSPQGWV
VRAAHLTPLLEYVPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSMRTQQAQ
QEAELTPRPAGVVPGA

445/550

FIGURE 445

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGAATGGACTGGCCTCACAACCTGC
TGTTTCTTCTTACCATTTCCATCTTCCTGGGGCTGGGCCAGCCCAGGAGCCCCAAAAGCAAGA
GGAAGGGGCAAGGGCGGCCTGGGCCCCCTGGCCCCCTGGCCCTCACCAGGTGCCACTGGACCTGG
TGTCACGGATGAAACCGTATGCCCCGATGGAGGAGTATGAGAGGAACATCGAGGAGATGGTGG
CCCAGCTGAGGAACAGCTCAGAGCTGGCCCAGAGAAAGTGTGAGGTCAACTTGCAGCTGTGGA
TGTCCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCCAGCCGTATCC
CCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGGCTGTGTGAACCCCTTCACCATGCAGG
AGGACCGCAGCATGGTGAGCGTGCCGGTGTTCAGCCAGGTTCCCTGTGCGCCGCCGCTCTGCC
CGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCGCTGTGGGCT
GCACCTGCATCTTCTGAATCACCTGGCCCAGAAGCCAGGCCAGCCCCGAGACCATCCTCCT
TGCACCTTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAAGCAAG

446/550

FIGURE 446

MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQVPLDLVSRMKPYARMEEYE
RNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPSRIPVDLPEARCLCLGC
VNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPPRTGPCRQRAVMETIAVGCTCIF

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

447/550

FIGURE 447

GGAGTGCAGATGGCATCCTTCGGTTCTTCCAGACAAGCTGCAAGACGCTGACCATGGCCAAGA
TGGAGCTCTCGAAGGCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCTAT
CACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCACACAGAAGGTGC
CCAAGCCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCCTGGATG
GAGATACCAACACATCCACCCAGGAGGTGGTACAATACTGGGAGACTGGGGATGACCGGT
TCTCCTTCCGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAACTGTGGAAGAACCAG
GGGAGAGGTGCCGAAGTTTCATTGAACTTACACCACCAGCCAAGAGAGGTGAGAAAGGACTAC
TGGAATTTGCCAGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAAGCGGTTGA
TGGAGAAGGCTTCCCTCCCCTCCCCTCCCTTGGGGCTTTGTGGCAAAAATCCTATGGTTATCC
CTGGGAACGCAGATCACCTACATCGGACTTCAATTCATCAGCTTCCTCCTGCTACTAACAGAC
TTGCTACTCACTGGGAACCCTGCCTGTGGGCTCAAAGTGAAGCGCCTTTGCTGCTGTTTCCTCT
GTCCTGTCAGGTCTCCTGGGGATGGTGGCCACATGATGTATTCAACAAGTCTTCCAAGCGACT
GTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGGAATTATGGCTGGGCCTTCTACATG
GCCTGGCTCTCCTTACCTGCTGCATGGCGTCGGCTGTCAACACCTTCAACACGTACACCAGG
ATGGTGCTGGAGTTCAAGTGCAAGCATAGTAAGAGCTTCAAGGAAAACCCGAAGTGCCTACCA
CATCACCATCAGTGTTTCCCTCGGCGGCTGTCAAGTGCAGCCCCCACCCTGGGTCTTTGACC
AGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGGGAGTCGACTTCTACTCC
GAGCTGCGGAACAAGGGATTTCAAAGAGGGGCCAGCCAGGAGCTGAAAGAAGCAGTTAGGTCA
TCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGAGTAGGCTTGAGCCCTACCT
TACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCGTCTCTTGAGCATGGTTTTTA
GAGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTCCTAAGGGATTCTGGGTGCCA
CTGCTCTCTTTTCTCTACAGCTCCATCTTGTTTACCCACCCACATCTCACACATCCAGAA
TTCCCTTCTTTACTGATAGTTTCTGTGCCAGGTTCTGGGCTAAACCATGGAGATAAAAAGAAG
AGTAAAATACACTTCCCGACCTTAAGGATCTGAAA

448/550

FIGURE 448

MAKMELSKAFSGQRTLLSAILSMLSLSFSTTSLLSNYWFVGTQKVPKPLCEKGLAAKCFDMPV
SLDGDNTNTSTQEVVQYNWETGDDRFSSFRSGMWLSCEETVEEPGERCRSFIELTPPAKRGE
KGLLEFATLQGPCHPTLRFGGKRLMEKASLPSPPLGLCGKNPMVIPGNADHLHRTSIHQLPPA
TNRLATHWEPCLWAQTERLCCCFLCPVRSPGDGGPHDVFTSLPSDCQLGSRRLTTCLELWLG
LLHGLALLHLLHGVGCHHLQHVVHQDGAGVQVQA

FIGURE 449

CCACGCGTCCGCCACGCGTCCGCCACGCGTCCGCCACGCGTCCGCCACGCGTCCGCC
ACGCGTCCGGTGCAAGCTCGCGCCGACACGTGCTGGTGGAGGAAAGGAGCCCGGGCGCTCTCGCCGCTCCCGC
CGCCGCGTCCGACACTCCCAACCGCCCGCGCGCCGCCGCCGCAAGCATGAGTGAGCCGCTCTCT
GCAGCTGCCCGGGGCGGAATGGCAGGCTGTTTCCGCGGAGTAAAAGGTGGCGCGGTCAGTGGTCGTTTCCAAT
GACGGACATTAACGAGCTGTAGATCTGGGGAGTCGCGAGCCCGCGAGTTTGGAGTTTCCCCCCCAACAGCT
CACAGTCCGAAGTCGAGAGGGGAAAGGAAGGCGCAGGAAGCGAAGCTCGGGCTCGGCACGTAGTTGGGAAACT
TGCGGGTCTAGAAAGTCGCTCCCCGCCTTGCCGGCCGCCCTTGACAGCCCCGAGCCGAGCAGCAAAGTGAGACAT
TGTGCGCTGCCAGATCCCGCGGGCGCGGCGGGGCTGCCTCGGAAACACAGAGGGGTCTTCTCGCCCTGCA
TATAATTAGCTGTGACACAAAGGGAGCAGCTGAATGGAGTTGTCTACTCTGAAAAAGGATTTCTGACCGAGCG
CTTCCCAATGGACATTTCCAGTCTCTCTGGAAAGATTCTCGTAATGGATTCTGCTGCTCGGTCTCTGTCTA
ACTGGCTGCTGAGGAGGCCCTCGGGGTGGTCTGTGTCTGCTGGGGCTGCTTTCAGATGCTGCCCGCGCCC
CCAGCGGTTGCCCGAGCTGTGCGGTCGAGAGGGCGGCTGCTGTACTCGAGGCGCTCAACCTCACAGAGGCGC
CCCCAACACTGTCCGGCTGCTGGGCTGTGCTCGCTACACAGCCTCTCGAGCTGCGCGCGCGCGCAGTTTCA
CGGGTTAATGCGACTCAGTGGCTCTATCTGATCAACATACATCTGCTCCGTGACGGGGAGCCCTTTAGA
AACTGCGCGGAGTTAAGGAACTCAGCTGAGTTTCAACACAGTACACCAACTGCCAACACCACCTTCGGGCCCA
TGCCCAACTCGCGCAGGCTGGACCTCTGTTACAAACAGCTCAGGCGCTCGCGCCGACCTCTTACCGGCTGC
GGAAGCTCACACGCTGCATATGCGGGCCACGCCATCCAGTTTGTGCGCGTGGCATCTTCCAGGACTCGCGCA
GCCTCAAGTTTCTCGACATCGGATACAATCAGCTCAAGAGTCTGGCGCGCAACTCTTTCGCCGGCTTGTTTAAGC
TACCAGAGCTGCACCTCGAGCAACAACGATTTGGTCAAGGTGAATCTGCCCACTTCGCCGCTCATCTCCCTGC
ACTCGCTCTGCTCGGAGGAACAAGGTGGCCATTTGTGTGCTGAGCTCGCTGGACTGGGTTTGAACCTGGAGAAAA
TGGAGTTGTGCGGGACAGAGATCGAGTACATGGAGCCCATGTGTTGCGAGACCTGCGCCACCTGCACTCCCTGC
AGCTGACTCCAACCGCTCACTACATCGAGCCCCGGATCTCAACTCTTGAAGTCCCTGACAAGCATCACCC
TGCGCGGAACTGTGGGATTGCGGGGCGCAAGCTGTGTGCTAGCCTCGTGCTCAGCAACTTCAGGGGCGCT
ACGATGGCAACTTGCAGTGCGCCAGCCCGGAGTACGCACAGGCGGAGGACGCTGTCGACCGCGTACGCTCTCC
ACCTGTGCGAGGATGGGGCCAGGCCACAGCGGCCACCTGCTCTGCGCCGTACCAACCGCAGTGATCTGGGG
CCCCTGCCAGCTCGGCCACCACGCTCGCGGACGGCGGGGAGGGGACGACGAGGCACATTGAGCCTGCCACCG
TGCTCTTCCAGGGCGCGAGCAGCCGGAAGACCGTGCAGATCCAAAGGTTGTCACGGGCACCATGGCCCTCTC
TCTTCTCTCTCTCATGCTGGTCTGCTGTCTGCTCTGAGTCTCGGAAGTGTTCACGACCGCTCAGGCAGCTCA
GACAGTGCTTTGTACGCGAGCGCAGGAAGCAAAAGCAGAAACAGACATGCATCAGATGGCTGCCATGCTTGCC
AGGAATACATACGTTGATTACAAACGAAACCATTTAGGGAGGCCCTGGTGATCATCAACAGGATATGGCTCGTGT
CTGCCACACGAGCGCCGCGAGGGAATCGAGGTGTAATGTGCCAGTGGCTCAACCCATCGCTACCAATA
CGCTCGGGCAGCGCGGACGGCGCGGCGGCCACGAGCTGGGCTCTCTTGTCTGTGCTCTGATATGCTCATTGAC
TGAACCTTTAAGGGATCTCTCCAGAGACTTGACATTTAGCTTTATTGTGCTTTAAAAACAAGGCAATTA
AACACAACAAAAACCCACCCACAACTTCAGGACAGTCTATCTTAAATTTTCATATGAGAACTCTCTCTCC
TTTGAAGATGCTTCATCTTACGAATCTGAGAGTGTAAAAAAGGTGGCCATAAGACAGAGAGGAATAATCGTG
CTTTGTTTTATGCTACTCTCTCCACCTGCGCATGATTAACATCATGTATGATAGAAGATCTTAAGTCCATACG
ATTTTCATGAAGAAGCAATTTGAAAGAGGAATCTGCAATCTGGGAGCTTAAGAGCAATGATGACCATAGAAGCT
TGTTCTTACTTTGTGTGTGCTGTATGTTTCTGCGTGTGTGTCTTTGTAGGCAAGCAACGTTGTCTACACA
AACGGGAATTTAGCTCACATATTTTCATGCCCCCTGTGCTCTAGCTCTGGAGATTTGGTGGGGGAGGTGGGGG
AACGGCAGGAATAAGGGAAGGTGGTATTTTAACTAAGTPTTTGTAACACTTGAATCTTTCTTTCTCAAAAT
ATTATCTTTAAGCTTCAAGAACTTGCTGACCCCTCTAAGCAACTACTAAGCATTTTAAAGAGAACTTAAT
TTTTAAGGTGTAGCACCTTTTTTTTTTATCTTCCACAGAGGGTGTATCTCATTATGCTGTGCTATCTGAAAA
GAACTTAAGGCCAATTAACGTTCTGCTCTGGGCTATGTGATGGATTGACCTTCAATTTGAGTACCTTCCAG
CTGATTAAGTTAGCAGAGTGGTATGAGTTTTCGAATATTATATAGAAAAAAGTCTTTTACATGACAAAT
GACACTCTCACACCACTTTAGCCCTAGTAGTTTTTTAGGTTGACACAGAGGAAGCAGGTTAAATGAGACCTGTC
CTCTGCTGCACTCGAAAAATAGGCAGTCCCTGATGCTCAGATCTTAGCCTTGATATTAATAGTTGAGACGAA
TACCCACAATGACAGCTATACTCCAGACTACAAAGTTACCATCGAAAGGAAGGTTATTCAGTAAAAAGGAA
ATAGTTTTCTCAACCATTTAAAAATATCTTCTGAATCATCAAAGTAGAAGAGCCCCCACTTTTCTCTGTC
CTTCAAGAAGGCAGACATTTGGTATGATTTAGCATCAACAACATTAATGAGTATGTAAGTATGAGAGGGG
CAAATGCCACTTGTATTCTCTCCAGTTTTCGAAGCAAGTACACACAGATCTCTGGTAGGATAGGGGCCACTT
GTGTTTCCGGCTTATTTTAGTCGACTGTGACGAAAGTTTGTAGCTGATCTATCTGACATGGCCAGTAGAACG
GGCATTAGGATGATACATGAGATGTTGAGGAAGCAATCATCATACACCTCTCACAGAGAAATATCAAGAA
CCAGAAATTTATCTGTTTTGGAGCAAGAGTGTCTAATGTTTCAGGGTAGTCAAATAAACATAAATATCTCC
TCTAGATGAGTGGCGATGTGGGCTATTTGGGCTGCGATTGACAGATGTCAAATAAAAAAGGAATTAGCTAGAA
TATGACCATTAATGTGCTTCTGAATATATTTTGGATAGGTTTAGAATGTCA

450/550

FIGURE 450

MDFLLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPHN
LSGLLGLSLRYNSLSELRAGQFTGLMQLTWLYLDHNNHICSVQGDFAFQLRRVKELTLSSNQIT
QLPNTTFRPMPNLRSDLSYNKLQALAPDLFHGLRKLTTLHMRANAIQFVPVRIFQDCRSKF
LDIGYNQLKSLARNSFAGLFKLTELHLEHNDLVKVNFAHFPRLLSLHSLCLRRNKVAIVVSSL
DWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRLTYIEPRILNSWKSLSITLAGNL
WDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDVYAFHLCEDGAEPSTGHLLSAV
TNRSDLGPPASSATTLADGGEGQHDGTFEPATVALPGGEHAENAVQIHKVVTGTMALIFSFLI
VVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHQMAAMSAQEYYVDYKPNHIEGALVII
NEYGSCTCHQQPARECEV

451/550

FIGURE 451

TTGAGCGCAGGTGAGCTCCTGCGCGTTCCGGGGGCGTTCCCTCCAGTCACCCTCCCGCCGTTAC
CCGCGGCGCGCCCCGAGGGAGTCTCCTCCAGACCTCCCTCCCGTTGCTCCAAACTAATACGGA
CTGAACGGATCGCTGCGAGGGTGGGAGAGAAAATTAGGGGGAGAAAGGACAGAGAGAGCAACT
ACCATCCATAGCCAGATAGATTATCTTACACTGAACTGATCAAGTACTTTGAAAATGACTTTCG
AAATTTATCTTGGTGTCTTCATACCTTGCTGCACTGAGTCTTTCAACCACCTTTTCTCTCCAA
CTAGACCAGCAAAAGGTTCTACTAGTTTCTTTTGATGGATTCCGTTGGGATTACTTATATAAA
GTTCCAACGCCCCATTTTCATTATATTATGAAATATGGTGTTCACGTGAAGCAAGTTACTAAT
GTTTTTATTACAAAACCTACCTAACCATTATACTTTGGTAACTGGCCTCTTTGCAGAGAAT
CATGGGATTGTTGCAAATGATATGTTTGATCCTATTTCGGAACAAATCTTTCTCCTTGGATCAC
ATGAATATTTATGATTCCAAGTTTGGGAAGAAGCGACACCAATATGGATCACAAACCAGAGG
GCAGGACATACTAGTGGTGCAGCCATGTGGCCCGGAACAGATGTAAAAATACATAAGCGCTTT
CCTACTCATTACATGCCTTACAATGAGTCAGTTTCATTTGAAGATAGAGTTGCCAAAATTGTT
GAATGGTTTACGTCAAAGAGCCCATAAATCTTGGTCTTCTCTATTGGGAAGACCTGATGAC
ATGGGCCACCATTTGGGACCTGACAGTCCGCTCATGGGGCCTGTCATTTAGATATTGACAAG
AAGTTAGGATATCTCATACAAATGCTGAAAAGGCAAGTTGTGGAACACTCTGAACCTAATC
ATCACAAGTGATCATGGAATGACGCAGTGCTCTGAGGAAAGGTTAATAGAAGTTGACCAGTAC
CTGGATAAAGACCCTATACCTGATTGATCAATCTCCAGTAGCAGCCATCTTGCCAAAAGAA
GGTAAATTTGATGAAGTCTATGAAGCACTAACTCACGCTCATCCTAATCTTACTGTTTACAAA
AAAGAAGACGTTCCAGAAAGGTGGCATTACAAATACAACAGTCGAATTCACCAATCATAGCA
GTGGCTGATGAAGGTGGCACATTTTACAGAATAAGTCAGATGACTTTCTGTTAGGCAACCAC
GGTTACGATAATGCGTTAGCAGATATGCATCCAATATTTTATAGCCCATGGTCTGCCTTCAGA
AAGAATTTCTCAAAGAAGCCATGAACTCCACAGATTTGTACCCACTACTATGCCACCTCCTC
AATATCACTGCCATGCCACACAATGGATCATTCTGGAATGTCCAGGATCTGCTCAATTCAGCA
ATGCCAAGGGTGGTCCCTTATACACAGAGTACTATACTCCTCCCTGGTAGTGTTAAACCAGCA
GAATATGACCAAGAGGGGTATACCCTTATTTTATAGGGGTCTCTCTTGGCAGCATTATAGTG
ATTGTATTTTTTGTAAATTTTCATTAAGCATTAAATTCACAGTCAAATACCTGCCTTACAAGAT
ATGCATGCTGAAATAGCTCAACCATTATTACAAGCCTAAATGTTACTTTGAAGTGGATTTGCAT
ATTGAAGTGGAGATTCCATAATTATGTCAGTGTTTAAAGGTTTCAAATTCTGGGAAACCAGTT
CCAAACATCTGCAGAAACCATTAAAGCAGTTACATATTTAGGTATACACACACACACACACA
CACATACACACACACGGACCAAAATACTTACACCTGCAAAGGAATAAAGATGTGAGAGTATGT
CTCCATTGTTCACTGTAGCATAGGGATAGATAAGATCCTGCTTTATTTGGACTTGGCGCAGAT
AATGTATATATTTAGCAACTTTGCACTATGTAAAGTACCTTATATATTGCACTTTAAATTTCT
CTCCTGATGGGTACTTTAATTTGAAATGCACTTTATGGACAGTTATGTCTTATAACTTGATTG
AAAATGACAACCTTTTTGCACCCATGTCACAGAATACTTGTACGCATTGTTCAAACCTGAAGGA
AATTTCTAATAATCCCGAATAATGAACATAGAAATCTATCTCCATAAATTGAGAGAAGAAGAA
GGTGATAAGTGTTGAAAATTAAATGTGATAACCTTTGAACCTTGAATTTTGGAGATGTATTCC
CAACAGCAGAATGCAACTGTGGGCATTTCTTGTCTTATTTCTTTCCAGAGAACGTGGTTTTCA
TTTATTTTTTCCCTCAAAGAGAGTCAAATACTGACAGATTGCTTCTAAATATATTGTTTCTGT
CATAAAATTATTTGTGATTTCTGATGAGTCATATTACTGTGATTTTCATAATAATGAAGACAC
CATGAATATACTTTTTCTTCTATATAGTTTCAAGCAATGGCCTGAATAGAAGCAACCAGGCACCAT
CTCAGCAATGTTTTCTTGTGTTGTAATTATTTGCTCCTTTGAAAATTAAATCACTATTAATT
ACATTAAAAATCAAATTGGATAAAAAAAAAAAAAAAAAAAAAA

452/550

FIGURE 452

MTSKFILVSFILAALSLSTTFSLQLDQQKVLLVSFDGFRWDYLYKVPTPHFHYIMKYGVHVKQ
VTNVFITKTPNHYTLVTGLFAENHGIVANDMFDPIRNKSFSLDHMNIYDSKFWEETPIWIT
NQRAGHTSGAAMWPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLYWED
PDDMGHHLGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEERLIEL
DQYLDKDHYTLLDQSPVAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYNRIQP
IIA VADEGWHILQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTDLYPLLC
HLLNITAMPHNGSFWNVQDLLNSAMPRVVPYTQSTILLPGSVKPAEYDQEGSYPIYFIGVSLGS
IIVIVFFVIFIKHLIHSQIPALQDMHAEIAQPLLQA

Important features:**Signal Peptide:**

amino acids 1-22

Transmembrane Domain:

amino acids 429-452

N-glycosylation sites:amino acids 101-104, 158-161, 292-295, 329-332, 362-365, 369-372,
382-385, 389-392**Somatomedin B Domain:**

amino acids 69-85

Sulfatase protein Region:

amino acids 212-241

453/550

FIGURE 453

GGCCGCCTGGAATTGTGGGAGTTGTGTCTGCCACTCGGCTGCCGGAGGCCGAAGGTCCGTGAC
T**ATG**GCTCCCCAGAGCCTGCCTTCATCTAGGATGGCTCCTCTGGGCATGCTGCTTGGGCTGCT
GATGGCCGCCTGCTTCACCTTCTGCCTCAGTCATCAGAACCTGAAGGAGTTTGGCCTGACCAA
CCCAGAGAAGAGCAGCACCAAAGAAACGGAGAGAAAAGAAACCAAAGCCGAGGAGGAGCTGGA
TGCCGAAGTCCTGGAGGTGTTCCACCCGACGCATGAGTGGCAGGCCCTTCAGCCAGGGCAGGC
TGTCCCTGCAGGATCCCACGTACGGCTGAATCTTCAGACTGGGGAAAGAGAGGCCAAAACCTCCA
ATATGAGGACAAGTTCCGAAATAATTTGAAAGGCCAAAAGGCTGGATATCAACACCAACACCTA
CACATCTCAGGATCTCAAGAGTGCCTGGAATAATTCAAGGAGGGGGCAGAGATGGAGAGTTT
AAAGGAAGACAAGGCAAGGCAGGCTGAGGTAAAGCGGCTCTTCCGCCCCATTGAGGAACTGAA
GAAAGACTTTGATGAGCTGAATGTTGTTCATTGAGACTGACATGCAGATCATGGTACGGCTGAT
CAACAAGTTCAATAGTTCCAGCTCCAGTTTGAAGAGAAGATTGCTGCGCTCTTTGATCTTGA
ATATTATGTCCATCAGATGGACAATGCGCAGGACCTGCTTTCCCTTGGTGGTCTTCAAGTGGT
GATCAATGGGCTGAACAGCACAGAGCCCCCTCGTGAAGGAGTATGCTGCGTTTGTGCTGGGCGC
TGCTTTTCCAGCAACCCCAAGGTCCAGGTGGAGGCCATCGAAGGGGGAGCCCTGCAGAAGCT
GCTGGTCATCCTGGCCACGGAGCAGCCGCTCACTGCAAAGAAGAAGGTCTGTTTGCCTGTG
CTCCCTGCTGCGCCACTTCCCCTATGCCAGCGGCAGTTCTGAAGCTCGGGGGGCTGCAGGT
CCTGAGGACCCTGGTGCAGGAGAAGGGCACGGAGGTGCTCGCCGTGCGCGTGGTCACACTGCT
CTACGACCTGGTCACGGAGAAGATGTTGCGCGAGGAGGAGGCTGAGCTGACCCAGGAGATGTC
CCCAGAGAAGCTGCAGCAGTATCGCCAGGTACACCTCCTGCCAGGCCTGTGGGAACAGGGCTG
GTGCGAGATCACGGCCCACCTCCTGGCGCTGCCCGAGCATGATGCCCCGTGAGAAGGTGCTGCA
GACACTGGGCGTCCTCCTGACCACCTGCCGGGACCGCTACCGTCAGGACCCCCAGCTCGGCAG
GACACTGGCCAGCCTGCAGGCTGAGTACCAGGTGCTGGCCAGCCTGGAGCTGCAGGATGGTGA
GGACGAGGGCTACTTCCAGGAGCTGCTGGGCTCTGTCAACAGCTTGCTGAAGGAGCTGAGAT**G**
AGGCCCCACACCAGGACTGGACTGGGATGCCGCTAGTGAGGCTGAGGGGTGCCAGCGTGGGTG
GGCTTCTCAGGCAGGAGGACATCTTGGCAGTGCTGGCTTGGCCATTAAATGGAAACCTGAAGG
CCAA
AA

454/550

FIGURE 454

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPESSTKETERKETKAEHEELD
AEVLEVFHPTHEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLKGKRLDINTNTY
TSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMVRLI
NKFSSSSSSLEEKIAALFDLEYYVHQMDNAQDLLSFGGLQVVINGLNSTEPLVKEYAAFLVGA
AFSSNPKVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKLGGGLQV
LRTLQVEKGTEVLAVRVVTLTYDLVTEKMFEEEEAELTQEMSPEKLQQYRQVHLLPGLWEQGW
CEITAHLLALPEHDAREKVLQTLGVLLTTCRDYRQDPQLGRTLASLQAEYQVLASLELQDGE
DEGYFQELLGSVNSLLKELR

Important features:**Signal peptide:**

amino acids 1-29

Hypothetical YJL126w/YLR351c/yhcX family protein.

amino acids 364-373

N-glycosylation site.

amino acids 193-197, 236-240

N-myristoylation site.

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

Homologous region SLS1 protein.

amino acids 68-340

455/550

FIGURE 455

GCCCCAGGGAGCAGTGGGTGGTTATAACTCAGGCCCGGTGCCAGAGCCCAGGAGGAGGCAGT
GGCCAGGAAGGCACAGGCCTGAGAAGTCTGCGGCTGAGCTGGGAGCAAATCCCCACCCCTA
CCTGGGGGACAGGGCAAAGTGAGACCTGGTGAGGGTGGCTCAGCAGGCAGGGAAGGAGAGGTGT
CTGTGCGTCCCTGCACCCACATCTTTCTCTGTCCCCTCCTTGCCCTGTCTGGAGGCTGCTAGAC
TCCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCGTCCTTG
TGTTTCCTCTCTACCTGGGGAAATAAGGTGCAGCGGCCATGGCTACAGCAAGACCCCCCTGGA
TGTGGGTGCTCTGTGCTCTGATCACAGCCTTGCTTCTGGGGGTACAGAGCATGTTCTCGCCA
ACAATGATGTTTCCTGTGACCACCCCTCTAACACCGTGCCCTCTGGGAGCAACCAGGACCTGG
GAGCTGGGGCCGGGGAAGACGCCCGGTGGATGACAGCAGCAGCCGCATCATCAATGGATCCG
ACTGCGATATGCACACCCAGCCGTGGCAGGCCGCGCTGTTGCTAAGGCCCAACCAGCTCTACT
GCGGGGCGGTGTTGGTGCATCCACAGTGGCTGCTCACGGCCGCCACTGCAGGAAGAAAGTTT
TCAGAGTCCGTCTCGGCCACTACTCCCTGTCAACAGTTTATGAATCTGGGCAGCAGATGTTCC
AGGGGGTCAAATCCATCCCCACCCTGGCTACTCCACCCTGGCCACTCTAACGACCTCATGC
TCATCAAACCTGAACAGAAGAATTCGTCCCACTAAAGATGTCAGACCCATCAACGTCTCCTCTC
ATTGTCCCTCTGCTGGGACAAAGTGCTTGGTGTCTGGCTGGGGGACAACCAAGAGCCCCCAAG
TGCACTTCCCTAAGGTCCTCCAGTGCTTGAATATCAGCGTGCTAAGTCAGAAAAGGTGCGAGG
ATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGACAAAGCAGGTAGAGACT
CCTGCCAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTCT
GGGAGATTACCCTTGTGCCCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTCA
CCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCATCCCAGGACTCAGCACACCGG
CATCCCCACCTGCTGCAGGGACAGCCCTGACACTCCTTTTCAGACCCTCATTCCTTCCCAGAGA
TGTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCCTGGACTCAGGGTCTGCTTCC
CCCACATTGGGCTGACCGTGTCTCTCTAGTTGAACCCTGGGAACAATTTCCAAAACCTGTCCAG
GGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCATCCTCAAGCTCAGGGCCCATCCCTT
CTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAATAAACTGAGAAGTGGAAGAAAAAAA

456/550

FIGURE 456

MATARPPMMWVLCALITALLGVTEHVLANNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDDS
SSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLSPV
YESGQQMFQGVKSIPHPGYSHPGHSNDLMLIKLNRRIRPTKDVRPINVSSHCP SAGTKCLVSG
WGTTKSPQVHF PKVLQCLNISVLSQKRCE DAYPRQIDDTMFCAGDKAGRDSCQGDSSGGPVVCN
GSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS

457/550

FIGURE 457

GCAGTCAGAGACTTCCCCTGCCCCCTGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCTTGCTTCCTGAACT
AGCTCACAGTAGCCCGGGCGGCCAGGGCAATCCGACCACATTTCACTCTCACCGCTGTAGGAATCCAGATGCAGG
CCAAGTACAGCAGCACGAGGGACATGCTGGATGATGATGGGGACACCACCATGAGCCTGCATTCTCAAGCCTCTG
CCACAACCTCGGCATCCAGAGCCCCGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGA
CCCTGCTGACTTTGTGCTTGGTGTCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTCAGTACTACCAGC
TCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATACGTCCCAAGAGTTGCAATCTC
TTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGCTGAAAACTCTGTCGTGAGCTGTATAACA
AAGCTGGAGCACACAGGTGCAGCCCTTGTACAGAACAATGGAAATGGCATGGAGACAATTGCTACCAGTTCTATA
AAGACAGCAAAAGTTGGGAGGACTGTAAATATTTCTGCCTTAGTGAAAACTCTACCATGCTGAAGATAAACAAAC
AAGAAGACCTGGAATTTGCCGCTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTTTGGCC
CTGACAGTGGCAAGGCCTGGCTGTGGATGGATGGAACCCCTTTCATTCTGAACTGTTCCATATTATAATAGATG
TCACCAGCCCCAAGAAGCAGAGACTGTGTGGCCATCTCAATGGGATGATCTTCTCAAAGGACTGCAAAGAATTGA
AGCGTTGTGTCTGTGAGAGAAGGGCAGGAATGGTGAAGCCAGAGAGCCTCCATGTCCCCCTGAAACATTAGGCG
AAGGTGACTGATTCGCCCTCTGCAACTACAAATAGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAG
ACATTGGGAAATGGAACATAATCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTCTCTCGTGTTCCTGTT
CAGGATCACCAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACAAGAAGTCTTATTTACATGCCAC
CAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTGGCTTAGAGATAACTTTTAGCTCTCTTTCTTCTCAA
TGTCTAATATCACCTCCCTGTTTTCATGTCTTCCCTTACACTTGGTGAATAAGAACTTTTGAAGTAGAGGAAA
TACATTGAGGTAACATCCTTTTCTCTGACAGTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTAC
CAGCAAATACACAAGGAATTCTTTTGTGTTGTTTCAGTTCATACTAGTCCCTTCCCAATCCATCAGTAAAGACCC
CATCTGCCTTGTCCATGCCGTTTCCCAACAGGGATGTCACTTGATATGAGAATCTCAAATCTCAATGCCTTATAA
GCATTCCCTCCTGTGTCCATTAAGACTCTGATAATTGTCTCCCTCCATAGGAATTTCTCCCAGGAAAGAAATAT
ATCCCCATCTCCGTTTCATATCAGAACTACCGTCCCCGATATCCCTTCAGAGAGATTAAAGACCAGAAAAAGT
GAGCCTCTTCATCTGCACCTGTAATAGTTTCAGTTCCATTTTTCTTCCATTGACCCATATTTATACCTTTCAGGT
ACTGAAGATTTAATAATAATAAATGTAAATACTGTGAAAA

458/550

FIGURE 458

MQAKYSSTRDMLDDDGDTTMSLHSQASATTRHPEPRRTEHRAPSSSTWRPVALTLLTLCLVLLI
GLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKLCRELY
NKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLEFAASQSY
SEFFYSYWTGLLRPDSGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAILNGMIFSKDCKEL
KRCVCERRAGMVKPESLHVPPETLGEGD

459/550

FIGURE 459

GTTGATGGCAAACCTTCCTCAAAGGAGGGGCAGAGCCTGCGCAGGGCAGGAGCAGCTGGCCCCAC
TGGCGGGCCCGCAACACTCCGTCTCACCTCTGGGCCCCACTGCATCTAGAGGAGGGCCGTCTGT
GAGGCCACTACCCCTCCAGCAACTGGGAGGTGGGACTGTCAGAAGCTGGCCCAGGGTGGTGGT
CAGCTGGGTCAGGGACCTACGGCACCTGCTGGACCACCTCGCCTTCTCCATCGAAGCAGGGAA
GTGGGAGCCTCGAGCCCTCGGGTGGAAGCTGACCCCAAGCCACCCTTACCTGGACAGGATGA
GAGTGTCAGGTGTGCTTCGCCCTCCTGGCCCTCATCTTTGCCATAGTCACGACATGGATGTTTA
TTCGAAGCTACATGAGCTTCAGCATGAAAACCATCCGTCTGCCACGCTGGCTGGCAGCCTCGC
CCACCAAGGAGATCCAGGTTAAAAAGTACAAGTGTGGCCTCATCAAGCCCTGCCAGCCAACT
ACTTTGCGTTTAAAATCTGCAGTGGGGCCGCCAACGTCGTGGGCCCTACTATGTGCTTTGAAG
ACCGCATGATCATGAGTCCTGTGAAAACAATGTGGGCAGAGGCCAAACATCGCCCTGGTGA
ATGGAACCACGGGAGCTGTGCTGGGACAGAAGGCATTTGACATGTACTCTGGAGATGTTATGC
ACCTAGTGAAATTCTTAAAGAAATTCGGGGGGGTGCACTGGTGCTGGTGGCCTCCTACGACG
ATCCAGGGACCAAAATGAACGATGAAAGCAGGAACTCTTCTCTGACTTGGGGAGTTCCTACG
CAAAACAACCTGGGCTTCGGGACAGCTGGGTCTTCATAGGAGCCAAAGACCTCAGGGGTAAAA
GCCCCTTTGAGCAGTTCTTAAAGAACAGCCAGACACAAACAAATACGAGGGATGGCCAGAGC
TGCTGGAGATGGAGGGCTGCATGCCCCGAAGCCATTTTAGGGTGGCTGTGGCTCTTCCTCAG
CCAGGGGCCTGAAGAAGCTCCTGCCTGACTTAGGAGTCAGAGCCCGGCAGGGGCTGAGGAGGA
GGAGCAGGGGGTGTGCGTGGAAGGTGCTGCAGGTCCTTGACGCTGTGTCGCGCCTCTCCTC
CTCGGAAACAGAACCCTCCACAGCACATCCTACCCGGAAGACCAGCCTCAGAGGGTCCTTCT
GGAACCAGCTGTCTGTGGAGAGAATGGGGTGCTTTCGTCAGGGACTGCTGACGGCTGGTCCTG
AGGAAGGACAACTGCCAGACTTGAGCCCAATTAAATTTTATTTTGTGCTGGTTTTGAAAAA
AAAAAAAAAAAAA

460/550

FIGURE 460

MRVSGVLRLLALIFAIVTTWMFIRSYMSFSMKTIRLPRWLAASPTKEIQVKKYKCGLIKPCPA
NYFAFKICSGAANVVGPTMCFEDRMIMSPVKNNVGRGLNIALVNGTTGAVLGQKAFDMYSGDV
MHLVKFLKEIPGGALVLVASYYDDPGTKMNDESRLKFLSDLGSSYAKQLGFRDSWVFIGAKDLRG
KSPFEQFLKNSPDTNKYEGWPELLEMEGCMPPKPF

Important features:

Signal peptide:

amino acids 1-15

ATP/GTP-binding site motif A (P-loop).

amino acids 184-191

N-glycosylation site.

amino acids 107-110

461/550

FIGURE 461

AAACTCAGCACTTGCCGGAGTGGCTCATTGTTAAGACAAAGGGTGTGCACTTCCTGGCCAGGA
AACCTGAGCGGTGAGACTCCCAGCTGCCTACATCAAGGCCCCAGGACATGCAGAACCTTCCTC
TAGAACCCGACCCACCACC**ATG**AGGTCCCTGCCTGTGGAGATGCAGGCACCTGAGCCAAGGCGT
CCAGTGGTCCTTGCTTCTGGCTGTCTTGGTCTTCTTTCTCTTCGCCTTGCCCTCTTTTATTAA
GGAGCCTCAAACAAAGCCTTCCAGGCATCAACGCACAGAGAACATTAAAGAAAGGTCTCTACA
GTCCCTGGCAAAGCCTAAGTCCCAGGCACCCACAAGGGCGAGGAGGACAACCATCTATGCAGA
GCCAGCGCCAGAGAAACAATGCCCTCAACACACAAACCCAGCCCAAGGCCACACCACCGGAGA
CAGAGGAAAGGAGGCCAACCAGGCACCGCCGGAGGAGCAGGACAAGGTGCCCCACACAGCACA
GAGGGCAGCATGGAAGAGCCCAAGAAAAGAGAAAACCATGGTGAACACACTGTCACCCAGAGG
GCAAGATGCAGGGATGGCCTCTGGCAGGACAGAGGCACAATCATGGAAGAGCCAGGACACAAA
GACGACCCAAAGGAAATGGGGGCCAGACCAGGAAGCTGACGGCCTCCAGGACGGTGTGAGAGAA
GCACCAGGGCAAAGCGGCAACCACAGCCAAGACGCTCATTTCCAAAAGTCAGCACAGAATGCT
GGCTCCCACAGGAGCAGTGTCAACAAGGACGAGACAGAAAGGAGTGACCACAGCAGTCATCCC
ACCTAAGGAGAAGAAACCTCAGGCCACCCACCCCTGCCCCTTTCCAGAGCCCCACGACGCA
GAGAAACCAAAGACTGAAGGCCGCCAACTTCAAATCTGAGCCTCGGTGGGATTTTGAGGAAAA
ATACAGCTTCGAAATAGGAGGCCTTCAGACGACTTGCCCTGACTCTGTGAAGATCAAAGCCTC
CAAGTCGCTGTGGCTCCAGAACTCTTTCTGCCCCAACCTCACTCTCTTCTGACTCCAGACA
CTTCAACCAGAGTGAGTGGGACCGCCTGGAACACTTTGCACCACCCTTTGGCTTCATGGAGCT
CAACTACTCCTTGGTGCAGAAAGTCTGTGACACGCTTCCCTCCAGTGCCCCAGCAGCAGTGCT
CCTGGCCAGCCTCCCCGCTGGGAGCCTCCGGTGCATCACCTGTGCCGTGGTGGGCAACGGGGG
CATCCTGAACAACCTCCCACATGGGCCAGGAGATAGACAGTCACGACTACGTGTTCCGATTGAG
CGGAGCTCTCATTAAGGCTACGAACAGGATGTGGGGACTCGGACATCCTTCTACGGCTTTAC
CGCCTTCTCCCTGACCCAGTCACTCCTTATATTGGGCAATCGGGGTTTCAAGAACGTGCCTCT
TGGGAAGGACGTCCGCTACTTGCACTTCTGGAAGGCACCCGGGACTATGAGTGGCTGGAAGC
ACTGCTTATGAATCAGACGGTGATGTCAAAAAACCTTTTCTGGTTCAGGCACAGACCCCAAGGA
AGCTTTTTCGGAAGCCCTGCACATGGACAGGTACCTGTTGCTGCACCCAGACTTTCTCCGATA
CATGAAGAACAGGTTTCTGAGGTCTAAGACCCTGGATGGTGCCCACTGGAGGATATACCGCCC
CACCCTGGGGCCCTCCTGCTGCTCACTGCCCTTCAGCTCTGTGACCAGGTGAGTGCTTATGG
CTTCATCACTGAGGGCCATGAGCGCTTTTCTGATCACTACTATGATACATCATGGAAGCGGCT
GATCTTTTACATAAACCATGACTTCAAGCTGGAGAGAGAAGTCTGGAAGCGGCTACACGATGA
AGGGATAATCCGGCTGTACCAGCGTCTGGTCCCAGCACTGCCAAAGCCAAGAAC**TGA**CCGGG
GCCAGGGCTGCCATGGTCTCCTTGCTGCTCCAAGGCACAGGATACAGTGCGGAATCTTGAGAC
TCTTTGGCCATTTCCCATGGCTCAGACTAAGCTCCAAGCCCTTCAGGAGTTCCAAGGGAACAC
TTGAACCATGGACAAGACTCTCTCAAGATGGCAAATGGCTAATTGAGGTTCTGAAGTTCTTCA
GTACATTGCTGTAGGTCTGAGGCCAGGGATTTTAAATTAATGGGGTGATGGGTGGCCAATA
CCACAATTCCCTGCTGAAAAACACTCTTCCAGTCCAAAAGCTTCTTGATACAGAAAAAAGAGCC
TGGATTTACAGAAACATATAGATCTGGTTTGAATTCAGATCGAGTTTACAGTTGTGAAATCT
TGAAGGTATTACTTAACTTCACTACAGATTGTCTAGAAGACCTTCTAGGAGTTATCTGATTC
TAGAAGGGTCTATACTTGTCTTGTCTTTAAGCTATTTGACAACCTCTACGTGTTGTAGAAAAAC
TGATAATAATACAAATGATTGTTGTCCATGGAAAGGCAAATAAATTTTCTACAGTGAAAAAA
AAAAAAA

462/550

FIGURE 462

MRSCLWRCRHLSQGVQWSLLLAVLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKPK
SQAPTRARRTTIYAEPAPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAWKS
PEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGNNGQTRKLTASRTVSEKHQGKAA
TTAKTLIPKSQHRMLAPTGA VSTRTRQKGVT TAVIPPKEKKPQATPPPAPFQSPTTQRNQRLK
AANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFLPNLTLFLDSRHFNQSEW
DRLEHFAPPFQFMELNYSLVQKVVT RFPVPQQQLLASLPAGSLRCITCAVVGNNGGILNNSH
MGQEIDSHDYVFRLSGALIKGYEQDVGTRTSFYGF TAFSLTQSLILGNRGFKNVPLGKD VRY
LHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEA FREALHMDRYLLHHPDFLR YMKNRFL
RSKTL DGAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYDTSWKRLIFYINH
DFKLEREVWKRLHDEGIIRLYQRP GPGTAKAKN

Important features:**Cytoplasmic Domain:**

amino acids 1-10

Type II Transmembrane Domain:

amino acids 11-35

Lumenal catalytic Domain:

amino acids 36-600

Ribonucleotide Reductase small subunit Signature:

amino acids 481-496

N-glycosylation Sites:

amino acids 300-303, 311-314, 331-334, 375-378, 460-463

463/550

FIGURE 463

GGGGGAGCTAGGCCGGCGGCAGTGGTGGTGGCGGGCGGCAAGGGTGAGGGCGGCCCCAGAAC
CCCAGGTAGGTAGAGCAAGAAGATGGTGTCTTCTGCCCCCAAATGGTCCCTTGCAACCATGTC
ATTTCTACTTTCTCACTGTTGGCTCTCTTAACGTGTGTCCACTCCTTCATGGTGTGACAGCAC
TGAAGCATCTCCAAAACGTAGTGATGGGACACCATTTCCTTGGAATAAAATACGACTTCCTGA
GTACGTCAATCCAGTTTCATTATGATCTCTTGATCCATGCAAACCTTACCACGCTGACCTTCTG
GGGAACCACGAAAGTAGAAATCACAGCCAGTCAGCCCACCAGCACCATCATCCTGCATAGTCA
CCACCTGCAGATATCTAGGGCCACCCTCAGGAAGGGAGCTGGAGAGAGGCTATCGGAAGAACC
CCTGCAGGTCCTGGAACACCCCCCTCAGGAGCAATTCGACTGCTGGCTCCCGAGCCCCCTCCT
TGTCGGGCTCCCGTACACAGTTGTCTATTCACTATGCTGGCAATCTTTCGGAGACTTTCCACGG
ATTTTACAAAAGCACCTACAGAACCAAGGAAGGGGAACCTGAGGATACTAGCATCAACACAATT
TGAACCCACTGCAGCTAGAATGGCCTTTCCTGCTTTGATGAACCTGCCTTCAAAGCAAGTTT
CTCAATCAAAATTAGAAGAGAGCCAAAGGCACCTAGCCATCTCCAATATGCCATTGGTGAAATC
TGTGACTGTTGCTGAAGGACTCATAGAAGACCATTTTGTATGTCAGTGTGAAGATGAGCACCTA
TCTGGTGGCCTTCATCATTTTCAGATTTTGGAGTCTGTGACGAAGATAACCAAGAGTGGAGTCAA
GGTTTCTGTTTATGCTGTGCCAGACAAGATAAATCAAGCAGATTATGCACTGGATGCTGCGGT
GACTCTTCTGAAATTTATGAGGATTATTTACGATACCGTATCCCTACCCAAACAAGATCT
TGCTGCTATTCCCGACTTTCAGTCTGGTGTATGGAATACTGGGGACTGACAACATATAGAGA
ATCTGCTCTGTTGTTTGTATGCAGAAAAGTCTTCTGCATCAAGTAAGCTTGGCATCACAGTGAC
TGTGGCCCATGAACTGGCCACCAGTGGTTTGGGAACCTGGTCACTATGGAATGGTGGAAATGA
TCTTTGGCTAAATGAAGGATTGCCAAATTTATGGAGTTTGTGTCTGTGAGTGTGACCCATCC
TGAAGTGAAGTTGGAGATTATTTCTTTGGCAATGTTTTGACGCAATGGAGGTAGATGGCTT
AAATTCCTCACACCCTGTGTCTACACCTGTGGAAAATCCTGCTCAGATCCGGGAGATGTTTGA
TGATGTTTCTTATGATAAGGGAGCTTGTATTCTGAATATGCTAAGGGAGTATCTTAGCGCTGA
CGCATTTAAAGTGGTATTGTACAGTATCTCCAGAAGCATAGCTATAAAAAATACAAAAACGA
GGACCTGTGGGATAGTATGGCAAGTATTTGCCCTACAGATGGTGTAAAAGGGATGGATGGCTT
TTGCTCTAGAAGTCAACATTATCTTCATCCTCACATTGGCATCAGGAAGGGGTGGATGTGAA
AACCATGATGAACACTTGGACACTGCAGAGGGGTTTTCCCTAATAACCATCACAGTGAGGGG
GAGGAATGTACACATGAAGCAAGAGCACTACATGAAGGGCTCTGACGGCGCCCCGGACACTGG
GTACCTGTGGCATGTTCCATTGACATTCATCACAGCAAAATCCAACATGGTCCATCGATTTT
GCTAAAAACAAAAACAGATGTGCTCATCCTCCAGAAAGAGGTGGAATGGATCAAATTTAATGT
GGGCATGAATGGCTATTACATTGTGCATTACGAGGATGATGGATGGGACTCTTTGACTGGCCT
TTTAAAGGAACACACACAGCAGTCAGCAGTAATGATCGGGCAAGTCTCATTAACAATGCATT
TCAGCTCGTCAGCATTTGGGAAGCTGTCCATTGAAAAGGCCTTGGATTTATCCCTGTACTTGAA
ACATGAAACTGAAATTATGCCCGTGTTCAGGTTTGAATGAGCTGATTCTATGTATAAGTT
AATGGAGAAAAGAGATATGAATGAAGTGGAAACTCAATTCAAGGCCTTCCTCATCAGGCTGCT
AAGGGACCTCATTTGATAAGCAGACATGGACAGACGAGGGCTCAGTCTCAGAGCAAAATGCTGCG
GAGTGAAGTACTACTCCTCGCCTGTGTGCACAACTATCAGCCGTGCGTACAGAGGGCAGAAGG
CTATTTTCAGAAAAGTGGAAAGGAATCCAATGGAAACTTGAGCCTGCCTGTGACGTGACCTTGGC
AGTGTGTTGCTGTGGGGGGCCAGAGCACAGAAGGCTGGGATTTTCTTTATAGTAAATATCAGTT
TTCTTTGTCCAGTACTGAGAAAAGCCAAATTGAATTTGCCCTCTGCAGAACCCAAAATAAGGA
AAAGCTTCAATGGCTACTAGATGAAAGCTTTAAGGGAGATAAAATAAAAACTCAGGAGTTTCC
ACAAATTTCTTACACTCATTGGCAGGAACCCAGTAGGATACCCACTGGCCTGGCAATTTCTGAG
GAAAACTGGAACAAACTTGTACAAAAGTTTGAACCTGGCTCATCTTCCATAGCCACATGGT
AATGGGTACAACAAATCAATTCTCCACAAGAACACGGCTTGAAGAGGTAAAAGGATTCTTCAG
CTCTTTGAAAGAAAATGGTTCTCAGCTCCGTTGTGTCCAACAGACAATTGAAACCATTGAAGA
AAACATCGGTTGGATGGATAAGAATTTTGATAAAATCAGAGTGTGGCTGCAAAGTGAAAAGCT
TGAACGTATGTAAAAATTCCTCCCTTGCCCGGTTCTCTGTTATCTCTAATCACCAACATTTTGT
TGAGTGTATTTTCAAAC TAGAGATGGCTGTTTTGGCTCCAACCTGGAGATACTTTTTTCCCTTC
AACTCATTTTTTGTACTATCCCTGTGAAAAGAATAGCTGTTAGTTTTTCATGAATGGGCTTTTT
CATGAATGGGCTATCGCTACCATGTGTTTTGTTTCATCACAGGTGTTGCCCTGCAACGTAAACC
CAAGTGTGGGTTCCCTGCCACAGAAGAATAAAGTACCTTATCTTCTCAAAAAAAAAAAAAA
AAAAAAAAAAAAA

464/550

FIGURE 464

MVFLPLKWSLATMSFLLSLLALLTVSTPSWCQSTEASPKRSDGTPFPWNKIRLPEYVIPVHY
DLLIHANLTTLTFWGTTKVEITASQPTSTIILHSHHLQISRATLRKGAGERLSEEPLQVLEHP
PQEQIALLAPELLVGLPYTVVIHYAGNLSETFHGFYKSTYRTKEGELRILASTQFEPTAARM
AFPCFDEPAFKASF SIKIRREPRHLAISNMPLVKSVTVAEGLIEDHFDVTVKMSTYLVAFIIS
DFESVSKITKSGVKVSVYAVPDKINQADYALDAAVTLLEFYEDYFSIPYPLPKQDLAAIPDFQ
SGAMENWGLTTYRESALLFDAEKSSASSKLGITVTVAHELAHQWFGNLVTMEWWNDLWLNELF
AKFMEFVS SVTHPELKVGDYFFGKCFDAMEVDALNSSHPVSTPVENPAQIREMFDDVSYDKG
ACILNMLREYLSADAFKSGIVQYLQKHSYKNTKNEDLWDSMASICPTDGVKGMDFGCSRSQHS
SSSSHWHQEGVDVKTMNTWTLQRGFPLITITVRGRNVHMKQEHYMKGSDGAPDTGYLWHVPL
TFITSKSNMVHRFLKTKTDVLILPEEVEWIKFNVGMNGYYIVHYEDDGWDSLTGLLKGTHTA
VSSNDRASLINNAFQLV SIGKLSIEKALDLSLYLKHETEIMPVFQGLNELIPMYKLMEKRD MN
EVETQFKAFLIRLLRDLIDKQTTWDEGSVSEQMLRSELLLLACVHNYQPCVQRAEGYFRKWKE
SNGNLSLPVDVTLAVFAVGAQSTEGWDFLYSKYQFSLSSSTEKSQIEFALCRTQNKELQWLLD
ESFKGDKIKTQEFQILTILGRNPVGYPLAWQFLRKNWNKLVQKFELGSSSIAHVMGTTNQF
STRTRLEEVKGFFSSSLKENGSQLRCVQQTIIETIEENIGWMDKNFDKIRVWLQSEKLERM

Important features:**Signal peptide:**

amino acids 1-34

N-glycosylation sites:

amino acids 70-74, 154-158, 414-418, 760-764, 901-905

Neutral zinc metallopeptidases, zinc-binding region signature:

amino acids 350-360

465/550

FIGURE 465

CAGCCACAGACGGGTCATGAGCGCGGTATTACTGCTGGCCCTCCTGGGGTTCATCCTCCCCT
GCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCCGA
CCTACCCCGGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGACAC
GTTGATGCTCATTGAGAGCGGACCCCAAGTGAGCCTGGTGCTCTCCAAGGGCTGCACGGAGGC
CAAGGACCAGGAGCCCCGCGTCACTGAGCACCGGATGGGCCCCGGCCTCTCCCTGATCTCCTA
CACCTTCGTGTGCCGCCAGGAGGACTTCTGCAACAACCTCGTTAACTCCCTCCCGCTTGGGC
CCCACAGCCCCCAGCAGACCCAGGATCCTTGAGGTGCCAGTCTGCTTGTCTATGGAAGGCTG
TCTGGAGGGGACAACAGAAGAGATCTGCCCAAGGGGACCACACACTGTTATGATGGCCTCCT
CAGGCTCAGGGGAGGAGGCATCTTCTCCAATCTGAGAGTCCAGGGATGCATGCCCCAGCCAGG
TTGCAACCTGCTCAATGGGACACAGGAAATTGGGCCCGTGGGTATGACTGAGAACTGCAATAG
GAAAGATTTTCTGACCTGTCATCGGGGGACCACCATTATGACACACGGAACTTGGCTCAAGA
ACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGGCAGGTGTGTGAGGAGAC
GCTGCTGCTCATAGATGTAGGACTCACATCAACCCTGGTGGGGACAAAAGGCTGCAGCACTGT
TGGGGCTCAAATTTCCAGAAGACCACCATCCACTCAGCCCCCTCCTGGGGTGCTTGTGGCCTC
CTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAGCAGCGTTCTGCTGAA
CTCCCTCCCTCCTCAAGCTGCCCCGTGCCAGGAGACCGGCAGTGTCTACCTGTGTGCAGCC
CCTTGGAACCTGTTCAAGTGGCTCCCCCGAATGACCTGCCCCAGGGGCGCCACTCATTGTTA
TGATGGGTACATTCATCTCTCAGGAGGTGGGCTGTCCACCAAAATGAGCATTCAGGGCTGCGT
GGCCCAACCTTCCAGCTTCTTGTGAACCACACCAGACAAATCGGGATCTTCTCTGCGCGTGA
GAAGCGTGATGTGCAGCCTCCTGCCTCTCAGCATGAGGGAGGTGGGGCTGAGGGCCTGGAGTC
TCTCACTTGGGGGGTGGGGCTGGCACTGGCCCCAGCGCTGTGGTGGGGAGTGGTTTGCCCTTC
CTGCTTAACTCTATTACCCCCACGATTCTTCACCGCTGCTGACCACCCACACTCAACCTCCCTC
TGACCTCATAACCTAATGGCCTTGGACACCAGATTCTTTCCCATTTCTGTCCATGAATCATCTT
CCCCACACACAATCATTCATATCTACTCACCTAACAGCAACACTGGGGAGAGCCTGGAGCATC
CGGACTTGCCCTATGGGAGAGGGGACGCTGGAGGAGTGGCTGCATGTATCTGATAATACAGAC
CCTGTCCTTTCA

466/550

FIGURE 466

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVKVSDLPRQWTPKNTSCDSGLGCQDTLMLIE
SGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQPPA
DPGSLRCPVCLSMEGCLEGTTEEICPKGTTTCYDGLLRRLRGGGIFSNLRVQGCMPOPGCNLLN
GTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSNTEMCEVGQVCQETLLID
VGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLNSLPPQ
AAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLSGGGLSTKMSIQGCVAQPSS
FLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLES LTWGVGLALAPALWWGVVCPSC

467/550

FIGURE 467

GAGGATTTGCCACAGCAGCGGATAGAGCAGGAGAGCACCACCGGAGCCCTTGAGACATCCTTG
AGAAGAGCCACAGCATAAGAGACTGCCCTGCTTGGTGTGTTTTGCAGGATGATGGTGGCCCTTCG
AGGAGCTTCTGCATTGCTGGTTCTGTTCCCTTGACGCTTTTCTGCCCCCGCCGAGTGTAACCA
GGACCCAGCCATGGTGCATTACATCTACCAGCGCTTTTCGAGTCTTGGAGCAAGGGCTGGAAAA
ATGTACCCAAGCAACGAGGGCATAACATTCAAGAATTCCAAGAGTTCTCAAAAAATATATCTGT
CATGCTGGGAAGATGTCAAGACCTACACAAGTGAGTACAAGAGTGAGTGGGTAACCTTGGCACT
GAGAGTTGAACGTGCCAACGGGAGATTGACTACATAACAATACCTTCGAGAGGCTGACGAGTG
CATCGTATCAGAGGACAAGACACTGGCAGAAATGTTGCTCCAAGAAGCTGAAGAAGAGAAAAA
GATCCGGACTCTGCTGAATGCAAGCTGTGACAACATGCTGATGGGCATAAAGTCTTTGAAAAAT
AGTGAAGAAGATGATGGACACACATGGCTCTTGGATGAAAGATGCTGTCTATAACTCTCCAAA
GGTGTACTTATTAATTGGATCCAGAAACAACACTGTTTGGGAATTTGCAAACATACGGGCATT
CATGGAGGATAACACCAAGCCAGCTCCCCGGAAGCAATCCTAACACTTTCCTGGCAGGGAAC
AGGCCAAGTGATCTACAAAGCTTTTCTATTTTTCATAACCAAGCAACTTCTAATGAGATAAT
CAAATATAACCTGCAGAAGAGGACTGTGGAAGATCGAATGCTGCTCCCAGGAGGGGTAGGCCG
AGCATTGGTTTACCAGCACTCCCCCTCAACTTACATTGACCTGGCTGTGGATGAGCATGGGCT
CTGGGCCATCCACTCTGGGCCAGGCACCCATAGCCATTTGGTTCTCACAAAGATTGAGCCGGG
CACACTGGGAGTGGAGCATTATGAGGATACCCCATGCAGAAAGCCAGGATGCTGAAGCCTCATT
CCTCTTGTGTGGGGTTCTCTATGTGGTCTACAGTACTGGGGGCCAGGGCCCTCATCGCATCAC
GTGCATCTATGATCCACTGGGCATATCAGTGAGGAGGACTTGCCCAACTTGTTCTTCCCCAA
GAGACCAAGAAGTCACTCCATGATCCATTACAACCCAGAGATAAGCAGCTCTATGCCTGGAA
TGAAGGAAACCAGATCATTTACAACTCCAGACAAAGAGAAAGCTGCCTCTGAAGTAATGCAT
TACAGCTGTGAGAAAGAGCACTGTGGCTTTGGCAGCTGTTCTACAGGACAGTGAGGCTATAGC
CCCTTCACAAATATAGTATCCCTCTAATCACACACAGGAAGAGTGTGTAGAAGTGGAATACGT
ATGCCTCCTTTCCCAATGTCAGTGCCTTAGGTATCTTCCAAGAGCTTAGATGAGAGCATATC
ATCAGGAAAGTTTCAACAATGTCCATTACTCCCCCAAACCTCCTGGCTCTCAAGGATGACCAC
ATTCTGATACAGCCTACTTCAAGCCTTTTGTGTTTACTGCTCCCCAGCATTACTGTAACCTCTG
CCATCTTCCCTCCACAAATTAGAGTTGTATGCCAGCCCCTAATATTACCACTGGCTTTTCTC
TCCCCTGGCCTTTGCTGAAGCTCTTCCCTCTTTTCAAATGTCTATTGATATTCTCCCATTTT
CACTGCCCCAACTAAAATACTATTAATATTTCTTTCTTTTCTTTTCTTTTCTTTTGGAGACAAGGT
CTCACTATGTTGCCAGGCTGGTCTCAAACCTCCAGAGCTCAAGAGATCCTCCTGCCTCAGCCT
CCTAAGTACCTGGGATTACAGGCATGTGCCACCACACCTGGCTTAAATACTATTTCTTATTG
AGGTTTAACTCTATTTCCCTAGCCCTGTCTTCCACTAAGCTTGGTAGATGTAATAATAAAA
GTGAAAATATTAACATTTGAATATCGCTTTCCAGGTGTGGAGTGTGTCACATCATTGAATTC
TCGTTTACCTTTGTGAACATGCACAAGTCTTTACAGCTGTCATTCTAGAGTTTAGGTGAGT
AACACAATTACAAAGTGAAAGATACAGCTAGAAAATACTACAAATCCCATAGTTTTTCCATTG
CCCAAGGAAGCATCAAATACGTATGTTTGTTCACCTACTCTTATAGTCAATGCGTTCATCGTT
TCAGCCTAAAAATAATAGTCTGTCCCTTTAGCCAGTTTTTCATGTCTGCACAAGACCTTTCAAT
AGGCCTTTCAAATGATAATTCCTCCAGAAAACAGTCTAAGGGTGAGGACCCCACTCTAGCC
TCCTCTTGTCTTGCTGTCTCTGTTTCTCTCTTTCTGCTTTAAATTCAATAAAAGTGACACTG
AGCAAAAAAAAAAAAAA

468/550

FIGURE 468

MMVALRGASALLVLFLAAFLPPPQCTQDPAMVHYIYQRFVLEQGLEKCTQATRAYIQEFQEF
SKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLLQE
AEEKKIRTLLNASCDNMLMGIKSLKIVKKMMDTHGSWMKDAVYNPKVYLLIGSRNNTVWEF
ANIRAFMEDNTKPAPRKQILTLWQGTGQVIYKGFLFFHNQATSNEIIKYNLQKRTVEDRMLL
PGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLTKEPGTLGVEHSWDTPCRSQ
DAEASFLLCGVLYVVYSTGGQGPHRITCIYDPLGTISEEDLPNLFFPKRPRSHSMIHYNPRDK
QLYAWNEGNQIIYKLQTKRKLPLK

469/550

FIGURE 469

TGGCCTCCCCAGCTTGCCAGGCACAAGGCTGAGCGGGAGGAAGCGAGAGGCATCTAAGCAGGC
AGTGTTTTGCCTTCACCCCAAGTGACCATGAGAGGTGCCACGCGAGTCTCAATCATGCTCCTC
CTAGTAACTGTGTCTGACTGTGCTGTGATCACAGGGGCTGTGAGCGGGATGTCCAGTGTGGG
GCAGGCACCTGCTGTGCCATCAGCCTGTGGCTTCGAGGGCTGCGGATGTGCACCCCGCTGGGG
CGGGAAGGCGAGGAGTGCCACCCCGGCAGCCACAAGGTCCCCTTCTTCAGGAAACGCAAGCAC
CACACCTGTCCTTGCTTGCCCAACCTGCTGTGCTCCAGGTTCCCGGACGGCAGGTACCGCTGC
TCCATGGACTTGAAGAACATCAATTTTTTAGGCGCTTGCTGGTCTCAGGATACCCACCATCCT
TTTCCTGAGCACAGCCTGGATTTTTTATTTCTGCCATGAAACCCAGCTCCCATGACTCTCCCAG
TCCCTACACTGACTACCCTGATCTCTCTGTCTAGTACGCACATATGCACACAGGCAGACATA
CCTCCCATCATGACATGGTCCCCAGGCTGGCCTGAGGATGTCACAGCTTGAGGCTGTGGTGTG
AAAGGTGGCCAGCCTGGTTCTCTTCCCTGCTCAGGCTGCCAGAGAGGTGGTAAATGGCAGAAA
GGACATTCCCCCTCCCCTCCCCAGGTGACCTGCTCTCTTTCCTGGGCCCTGCCCCTCTCCCA
CATGTATCCCTCGGTCTGAATTAGACATTCTTGGGCACAGGCTCTTGGGTGCATTGCTCAGAG
TCCCAGGTCTTGGCCTGACCCCTCAGGCCCTTCACGTGAGGTCTGTGAGGACCAATTTGTGGGT
AGTTCATCTTCCCTCGATTGGTTAACTCCTTAGTTTCAGACCACAGACTCAAGATTGGCTCTT
CCCAGAGGGCAGCAGACAGTCACCCCAAGGCAGGTGTAGGGAGCCCAGGGAGGCCAATCAGCC
CCCTGAAGACTCTGGTCCCAGTCAGCCTGTGGCTTGTGGCCTGTGACCTGTGACCTTCTGCCA
GAATTGTCATGCCTCTGAGGCCCCCTCTTACCACACTTTACCAGTTAACCCTGAAGCCCCCA
ATTCCCACAGCTTTTCCATTAAAATGCAAATGGTGGTGGTTCAATCTAATCTGATATTGACAT
ATTAGAAGGCAATTAGGGTGTTCCTTAAACAACCTCTTTCCAAGGATCAGCCCTGAGAGCAG
GTTGGTGACTTTGAGGAGGGCAGTCCTCTGTCCAGATTGGGGTGGGAGCAAGGGACAGGGAGC
AGGGCAGGGGCTGAAAGGGGCACTGATTGAGACCAGGGAGGCAACTACACACCAACATGCTGG
CTTTAGAATAAAAGCACCAACTGAAAAA

470/550

FIGURE 470

MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCAISLWLRGLRMCTPLGREGEETCHPG
SHKVPFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF

Important feratures:

Signal peptide:

amino acids 1-19

Tyrosine kinase phosphorylation site:

amino acids 88-95

N-myristoylation sites:

amino acids 33-39, 35-41, 46-52

471/550

FIGURE 471

AGCGCCCGGGCGTCGGGGCGGTAAAAGGCCGGCAGAAGGGAGGCACTTGAGAAATGTCTTTCC
TCCAGGACCCAAGTTTCTTCACCATGGGGATGTGGTCCATTGGTGCAGGAGCCCTGGGGGCTG
CTGCCTTGGCATTGCTGCTTGCCAACACAGACGTGTTTCTGTCCAAGCCCCAGAAAGCGGCCC
TGGAGTACCTGGAGGATATAGACCTGAAAACACTGGAGAAGGAACCAAGGACTTTCAAAGCAA
AGGAGCTATGGGAAAAAAATGGAGCTGTGATTATGGCCGTGCGGAGGCCAGGCTGTTTCCTCT
GTCGAGAGGAAGCTGCGGATCTGTCCTCCCTGAAAAGCATGTTGGACCAGCTGGGCGTCCCCC
TCTATGCAGTGGTAAAGGAGCACATCAGGACTGAAGTGAAGGATTTCCAGCCTTATTTCAAAG
GAGAAATCTTCCTGGATGAAAAGAAAAGTTCTATGGTCCACAAAGGCGGAAGATGATGTTTA
TGGGATTTATCCGTCTGGGAGTGTGGTACAACCTTCTCCGAGCCTGGAACGGAGGCTTCTCTG
GAAACCTGGAAGGAGAAGGCTTCATCCTTGGGGGAGTTTTCGTGGTGGGATCAGGAAAGCAGG
GCATTCTTCTTGAGCACCGAGAAAAAGAATTTGGAGACAAAGTAAACCTACTTTCTGTTCTGG
AAGCTGCTAAGATGATCAAACCACAGACTTTGGCCTCAGAGAAAAATGATTGTGTGAAACTG
CCCAGCTCAGGGATAACCAGGGACATTACCTGTGTTTCATGGGATGTATTGTTTCCACTCGTG
TCCCTAAGGAGTGAGAAACCCATTTATACTCTACTCTCAGTATGGATTATTAATGTATTTTAA
TATTCTGTTTAGGCCCCACTAAGGCAAAATAGCCCCAAAACAAGACTGACAAAAATCTGAAAAA
CTAATGAGGATTATTAAGCTAAAACCTGGGAAATAGGAGGCTTAAAATTGACTGCCAGGCTGG
GTGCAGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGAGCAAGTCACTTGAG
GTCGGGAGTTGAGACCAGCCTGAGCAACATGGCGAAACCCCGTCTCTACTAAAAATACAAAA
ATCACCCGGGTGTGGTGGCAGGCACCTGTAGTCCCAGCTACCCGGGAGGCTGAGGCAGGAGAA
TCACTTGAACCTGGGAGGTGGAGGTTGCGGTGAGCTGAGATCACACCACTGTATTCCAGCCTG
GGTGACTGAGACTCTAACTAA

472/550

FIGURE 472

MSFLQDPSFFTGMWSIGAGALGAAALALLANTDVFLSKPQKALEYLEDIDLKTLEKEPRT
FKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDFQP
YFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGWYNFFRAWNGGFSGNLEGEFGFILGGVFVVG
GKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

473/550

FIGURE 473

AATATATCATCTATTTATCATTAATCAATAATGTATTCTTTTATTCCAATAACATTTGGGTTT
TGGGATTTTAATTTTCAAACACAGCAGAATGACATTTTTTCTGTCACTATTATTATTGTTGGT
ATGTGAAGCTATTTGGAGATCCAATTCAGGAAGCAACACATTGGAGAATGGCTACTTTCTATC
AAGAAATAAAGAGAACCACAGTCAACCCACACAATCATCTTTAGAAGACAGTGTGACTCCTAC
CAAAGCTGTCAAACACAGGCAAGGGCATAGTTAAAGGACGGAATCTTGACTCAAGAGGGTT
AATTCTTGGTGCTGAAGCCTGGGGCAGGGGTGTAAAGAAAAACACTTAGATTCAATGATTGTA
AATTTAAGGCAAATACACATATTAGTATTACCTTAGTGTAATGTATCCCTGTCATATATACAA
TAAGGTGAAATTATAAGTACCCTATGCAGTTGGCTGGACAGTTCTAAATTGGACTTTATTAAT
TTTTAAATCAGTAACTGATTTATCACTGGCTATGTGCTTAGATCTACAGGAGATCATATAAT
TTGATACAAATAAAAGAAAAGTGTCTCTCCCCTTACAGAATTGACATTTTAAATGCGATACA
GTTAGAATAGGAAATATGACATTAGAAAGGAAGAATGACAGGGAGAAAGGAAAGAAGGGAAAA
TGTTGCCAAGGAAAAAAAAA

474/550

FIGURE 474

MTFFLSLLLLLVCEAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGKG
IVKGRNLDSRGLILGAEAWGRGVKKNT

476/550

FIGURE 476

MAPLALHLLVLPILLSLVASQDWKAERSQDPFEKCMQDPDYEQLLKVVVTWGLNRTLKPQRVI
VVGAGVAGLVAAKVLSDAGHKVTILEADNRIGGRIFTYRDQNTGWIGELGAMRMPSSHRILHK
LCQGLGLNLTKFTQYDKNTWTEVHEVKLRNYVVEKVPEKLGALRPQEKGHSPEDIYQMALNQ
ALKDLKALGCRKAMKKFERHTLLEYLLGEGNLSRPAVQLLDVMSSEGGFFYLSFAEALRAHSC
LSDRLQYSRIVGGWDLPRALLSSLSGLVLLNAPVVAMTQGPHDVHVQIETSPPARNLKVLKA
DVLLTASGPAVKRITFSPPLPRHMQEALRRLHYVPATKVFLSFRRPFWREEHIEGGSNTDR
PSRMIFYPFPPREGALLASYTWSDAFAAGLSREEALRLALDDVAALHGPVVRQLWDGTGVV
KRWAEDQHSQGGFVVQPPALWQTEKDDWTVPYGRIYFAGEHTAYPHGWVETAVKSALRAAIKI
NSRKGPASDTASPEGHASDMEGQGHVHGVASSPSHDLAKEEGSHPPVQGQLSLQNTTHTRTSH

Important features:**Signal peptide:**

amino acids 1-21

477/550

FIGURE 477

CTGACATGGCCTGACTCGGGACAGCTCAGAGCAGGGCAGAACTGGGGACACTCTGGGCCGGCCTTCTGCCTGCAT
GGACGCTCTGAAGCCACCCCTGTCTCTGGAGGAACCACGAGCGAGGGGAAGAAGGACAGGGACTCGTGTGGCAGGAA
GAACCTCAGAGCCGGGAAGCCCCATTCACTAGAAGCACTGAGAGATGCGGCCCCCTCGCAGGGTCTGAATTTCCCT
GCTGCTGTTCAAAAGATGCTTTTATCTTTAACTTTTTGTTTTCCCACTTCCGACCCCGCGTTGATCTGCAT
CCTGACATTTGGAGCTGCCATCTTCTGTGGCTGATCACCAGACCTCAACCCGTCTTACCTCTTCTTGACCTGAA
CAATCAGTCTGTGGGAATTGAGGGAGGAGCACGGAAGGGGGTTTTCCAGAGAACAATGACCTAACAAGTTGCTG
CTTCTCAGATGCCAAGACTATGTATGAGGTTTTCCAAAGAGGACTCGCTGTGTCTGACAATGGGCCCTGCTTGGG
ATATAGAAAACCAACAGCCCTACAGATGGCTATCTTACAAACAGGTGTCTGATAGAGCAGAGTACCTGGGTTT
CTGTCTCTTGCCATAAAGGTTATAAATCATCACCAGACCAAGTTTGTGCGCATCTTGTCTCAGAATAGGCCAGAGTG
GATCATCTCCGAATTGGCTTGTACACGTACTCTATGGTAGCTGTACCTCTGTATGACACCTTGGGACCAGAAGC
CATCGTACATATTGTCAACAAGCTGATATCGCCATGGTGATCTGTGACACACCCCAAAAGGCATTGGTGCTGAT
AGGGAATGTAGAGAAAGGCTTCAACCCGAGCCTGAAGGTGATCATCCTTATGGACCCCTTTGATGATGACCTGAA
GCAAAGAGGGGAGAAGAGTGAATGAGATCTTATCCCTATATGATGCTGAGAACCTAGGCAAAGAGCACTTCAG
AAAACCTGTGCCTCCTAGCCCAAGAGACCTGAGCGTCACTGTCTTACCAGTGGGACCACAGGTGACCCCAAGG
AGCCATGATAACCCATCAAAATATTGTTTTCAATGCTGCTGCCTTTCTCAAATGTGTGGAGCATGCTTATGAGCC
CACTCCTGATGATGTGGCCATATCCTACCTCCCTCTGGCTCATATGTTTGAGAGGATTGTACAGGCTGTTGTGTA
CAGCTGTGGAGCCAGAGTTGGATTCTTCCAAGGGGATATTCGGTTGCTGGCTGACGACATGAAGACTTTGAAGCC
CACATTGTTTCCCGCGGTGCTCGACTCCTTAACAGGATCTACGATAAGGTACAAAATGAGGCCAAGACACCCTT
GAAGAAGTTCTTGTGTAAGCTGGCTGTTTCCAGTAAATTCAAAGAGCTTCAAAGGGTATCATCAGGCATGATAG
TTTTCTGGGACAAGCTCATCTTTGCAAAGATCCAGGACAGCCTGGGCGGAAGGGTTCGTGTAATTGTCACTGGAGC
TGCCCCCATGTCCACTTCAGTCATGACATTCTTCCGGGCAGCAATGGGATGTCAAGGTGTATGAAGCTTATGGTCA
AACAGAATGCACAGGTGGCTGTACATTTACATTACCTGGGGACTGGACATCAGGTCACGTTGGGGTGCCCTGGC
TTGCAATTACGTGAAGCTGGAAGATGTGGCTGACATGAACACTTTACAGTGAATAATGAAGGAGAGGTCTGCAT
CAAGGGTACAAACGTGTTCAAAGGATACCTGAAGGACCCTGAGAAGACACAGGAAGCCCTGGACAGTGATGGCTG
GCTTCACACAGGAGACATTGGTCGCTGGCTCCCGAATGGAACCTCTGAAGATCATCGACCGTAAAAAGAACATTTT
CAAGCTGGCCCAAGGAGAATACATTGCACCAGAGAAGATAGAAAAATCTACAACAGGAGTCAACCAGTGTACA
AATTTTTGTACACGGGGAGAGCTTACGGTCATCCTTAGTAGGAGTGGTGGTTCCCTGACACAGATGTACTTCCCTC
ATTTGCAGCCAAGCTTGGGGTGAAGGGCTCCTTTGAGGAACCTGTGCCAAAACCAAGTTGTAAGGGAAGCCATTTT
AGAAGACTTGCAAAAATTTGGGAAAGAAAGTGGCCTTAAACTTTTGAAACAGGTCAAAGCCATTTTTCTTCATCC
AGAGCCATTTTCCATTGAAAATGGGCTCTTGACACCAACATTGAAAGCAAAGCGAGGAGAGCTTTCCAAATACCT
TCGGACCCAAATTGACAGCCTGTATGAGCACATCCAGGATTAGGATAAGGTACTTAAGTACCTGCCGGCCCACTG
TGCACTGCTTGTGAGAAAATGGATTAAAAACTATTCTTACATTTGTTTTGCCTTTCCCTCCTATTTTTTTTAACC
TGTTAAACTCTAAAGCCATAGCTTTTGTGTTTTATTTGAGACATATAATGTGTAACTTAGTTCCCAAATAAATCA
ATCCTGTCTTTCCCATCTTCGATGTTGCTAATATTAAGGCTTCAGGGCTACTTTTATCAACATGCCTGTCTTCAA
GATCCAGTTTATGTTCTGTGTCCTTCCCTCATGATTTCCAACCTTAATACTATTAGTAACCACAAGTTCAAGGGT
CAAAGGGACCCCTCTGTGCTTCTTCTTTGTTTTGTGATAAACATAACTTGCCAACAGTCTCTATGCTTATTTACA
TCTTCTACTGTTCAAACCTAAGAGATTTTAAATCTGAAAAACTGCTTACAATTCATGTTTTCTAGCCACTCCAC
AAACCACTAAATTTTAGTTTATGCTTATCACTCATGTCAATCATATCTATGAGACAAATGTCTCCGATGCTCTT
CTGCGTAAATTAATTTGTGTAAGGGAAGTGTGATCATACCAACATTTCCCTAACTCTCTAGTTAGATA
TCTGACTTGGGAGTATTAAAAATTGGGTCTATGACATACTGTCCAAAAGGAATGCTGTTCTTAAAGCATTATTTA
CAGTAGGAACCTGGGGAGTAAATCTGTTCCCTACAGTTTGTGCTGAGCTGGAAGCTGTGGGGGAAGGAGTTGACA
GGTGGGCCAGTGAACCTTTCCAGTAAATGAAGCAAGCACTGAATAAAAACCTCCTGAACTGGGAACAAAGATCT
ACAGGCAAGCAAGATGCCACACAACAGGCTTATTTTCTGTGAAGGAACCAACTGATCTCCCCACCCCTTGATT
AGAGTTCCTGCTCTACCTTACCCACAGATAACACATGTTGTTTCTACTTGTAATGTAAAGTCTTTAAATAAAC
TATTACAGATAAAAAA

478/550

FIGURE 478

MDALKPPCLWRNHERGKKDRDSCGRKNSEPGSPHSLEALRDAAPSQGLNFKLLFTKMLFIFNF
LFSPLPTPALICILTFGAAIFLWLITRPQPVLPDLLNNSVGIEGGARKGVSQKNNDLTSCC
FSDAKTMYEVFQRLAVSDNGPCLGYRKPNQPYRWLSYKQVSDRAEYLGSCLLHKGYSKSSPDQ
FVGIFAQNRPEWIISELACYTYSMAVAVPLYDTLGPEAIVHIVNKADIAMVICDTPQKALVLIG
NVEKGFTPSLKVIIIMDPFDDDLKQGEKSGIEILSLYDAENLGKEHFRKPVPPSPEDLSVIC
FTSGTTGDPKGAMITHQNIIVSNAAFLKCVEHAYEPTPDDVAISYLPLAHMFERIVQAVVYSC
GARVGFFQGDIRLLADDMKTLKPTLFPAVPRLLNRIYDKVQNEAKTPLKKFLLKLAIVSSKFKE
LQKGIIRHDSFWDKLIFAKIQDSLGRVRVIVTGAAPMSTSVMTFFRAAMGCQVYEAYGQTEC
TGGCTFTLPGDWTSGHVGVPACNYVKLEDVADMNYFTVNNEGEVCIKGTNVFKGYLKDPEKT
QEALDSGWLHTGDIGRWLPNGTLKIIDRKKNIFKLAQGEYIAPEKIENIYNRSQPVQLQIFVH
GESLRSSLVGVVVPDTPDLPSFAAKLGVKGSFEELCQNVVREAILEDLQKIGKESGLKTFEQ
VKAIFLHPEPFPSIENGLLTPTLKAKRGELSKYFRTQIDSLYEHIQD

Important features:**Type II transmembrane domain:**

amino acids 61-80

Putative AMP-binding domain signature.

amino acids 314-325

N-glycosylation site.

amino acids 102-105, 588-591 and 619-622

479/550

FIGURE 479

GGAGGCGGAGGCCGCGGCGAGCCGGGCGGAGCAGTGAGGGGCCCTAGCGGGGCCCGAGCGGGGC
CCGGGGCCCCCTAAGCCATTCTGAAGTCATGGGCTGGCCAGGACATTGGTGACCCGCCAATCC
GGTATGGACGACTGGAAGCCCAGCCCCCTCATCAAGCCCTTTGGGGCTCGGAAGAAGCGGAGC
TGGTACCTTACCTGGAAGTATAAACTGACAAACCAGCGGGCCCTGCGGAGATTCTGTGACACA
GGGGCCGTGCTTTTCTGCTGGTGACTGTCAATTGTCAATATCAAGTTGATCCTGGACACTCGG
CGAGCCATCAGTGAAGCCAATGAAGACCCAGAGCCAGAGCAAGACTATGATGAGGCCCTAGGC
CGCCTGGAGCCCCACGGGCGAGAGGCAGTGGTCCCCGGCGGGTCTTGACGTAGAGGTGTAT
TCAAGTCGCAGCAAAGTATATGTGGCAGTGGATGGCACCACGGTGCTGGAGGATGAGGCCCGG
GAGCAGGGCCGGGGCATCCATGTCATTGTCCTCAACCAGGCCACGGGCCACGTGATGGCAAAA
CGTGTGTTTGACACGTA CTACCTCATGAGGATGAGGCCATGGTGCTATTCCTCAACATGGTA
GCGCCCCGGCCGAGTGCTCATCTGCACTGTCAAGGATGAGGGCTCCTTCCACCTCAAGGACACA
GCCAAGGCTCTGCTGAGGAGCCTGGGCAGCCAGGCTGGCCCTGCCCTGGGCTGGAGGGACACA
TGGGCCTTCGTGGGACGAAAAGGAGGTCTGTCTTCGGGGAGAAACATTCTAAGTCACCTGCC
CTCTCTTCTGGGGGGACCCAGTCTCTGCTGAAGACAGATGTGCCATTGAGCTCAGCAGAAGAG
GCAGAGTGCCACTGGGCAGACACAGAGCTGAACCGTCGCCGCCGGCGCTTCTGCAGCAAAGTT
GAGGGCTATGGAAGTGATGCAGCTGCAAGGACCCACACCCATCGAGTTCAGCCCTGACCCA
CTCCAGACAACAAGTCTCAATGTGCCTGTGGCTGTCAATTGCAGGGAACCGACCCAATTAC
CTGTACAGGATGCTGCGCTCTCTGCTTTCAGCCCAGGGGTGTCTCCTCAGATGATAACAGTT
TTCATTGACGGCTACTATGAGGAACCCATGGATGTGGTGGCACTGTTTGGTCTGAGGGGCATC
CAGCATACTCCCATCAGCATCAAGAATGCCCGCGTGTCTCAGCACTACAAGGCCAGCCTCACT
GCCACTTTCAACCTGTTTCCGGAGGCCAAGTTTGTGTGGTTCTGGAAGAGGACCTGGACATT
GCTGTGGATTTTTTTCAGTTTCTGAGCCAATCCATCCACCTACTGGAGGAGGATGACAGCCTG
TACTGCATCTCTGCCTGGAATGACCAGGGGTATGAACACACGGCTGAGGACCCAGCACTACTG
TACCGTGTGGAGACCATGCCTGGGCTGGGCTGGGTGCTCAGGAGGTCCTTGTAACAAGGAGGAG
CTTGAGCCCAAGTGGCCTACACCGGAAAAGCTCTGGGATTGGGACATGTGGATGCGGATGCCT
GAACAACGCCGGGGCCGAGAGTGCATCATCCCTGACGTTTCCCGATCCTACCACTTTGGCATC
GTCGGCCTCAACATGAATGGCTACTTTACGAGGCCTACTTCAAGAAGCACAAGTTCAACACG
GTTCCAGGTGTCCAGCTCAGGAATGTGGACAGTCTGAAGAAAGAAGCTTATGAAGTGGAAGTT
CACAGGCTGCTCAGTGAGGCTGAGGTTCTGGACCACAGCAAGAACCCTTGTGAAGACTCTTTC
CTGCCAGACACAGAGGGCCACACCTACGTGGCCTTTATTTCGAATGGAGAAAGATGATGACTTC
ACCACCTGGACCCAGCTTGCCAAGTGCCTCCATATCTGGGACCTGGATGTGCGTGGCAACCAT
CGGGGCCTGTGGAGATTGTTTCGGAAGAAGAACCCTTCTGCTGGTGGTGGGGGTCCCGGCTTCC
CCCTACTCAGTGAAGAAGCCACCCTCAGTCACCCCAATTTTCTGGAGCCACCCCCAAAGGAG
GAGGGAGCCCCAGGAGCCCCAGAACAGACATGAGACCTCCTCCAGGACCCTGCGGGGCTGGGT
ACTGTGTACCCCCAGGCTGGCTAGCCCTTCCCTCCATCCTGTAGGATTTTGTAGATGCTGGTA
GGGGCTGGGGCTACCTTGTTTTAACATGAGACTTAATTACTAACTCCAAGGGGAGGGTTCCC
CTGCTCCAACACCCCGTTCTGAGTTAAAGTCTATTTATTTACTTCTTGTGGAGAAGGGC
AGGAGAGTACCTGGGAATCATTACGATCCCTAGCAGCTCATCCTGCCCTTTGAATACCCTCAC
TTTCCAGGCCTGGCTCAGAATCTAACCTATTTATTGACTGTCCTGAGGGCCTTGAAAACAGGC
CGAACCTGGAGGGCCTGGATTTCTTTTGGGCTGGAATGCTGCCCTGAGGGTGGGGCTGGCTC
TTACTCAGGAACTGCTGTGCCCAACCCATGGACAGGCCCAGCTGGGGCCACATGCTGACAC
AGACTCACTCAGAGACCCTTAGACACTGGACCAGGCCTCCTCTCAGCCTTCTCTTGTCCAGA
TTTCCAAAGCTGGATAAGTTGGTCATTGATTAAAAAAGGAGAAGCCCTCTGGGAAAAAAAAA
AAAAAAAAAAAAAAAAA

480/550

FIGURE 480

MDDWKPSPLIKPFGARKKRSWYLTWKYKLTNQALRRFCQTGAVLFLLVTIVNIKLILDTRR
AISEANEDPEPEQDYDEALGRLEPPRRRGSGPRRVLDVEVYSSRSKVYVAVDGTTVLEDEARE
QGRGIHVIVLNQATGHVMAKRVEDTYSPEHEDEAMVLFNLMVAPGRVLICTVKDEGSFHLKDTA
KALLRSLGSQAGPALGWRDTWAFVGRKGGPVFGEKHSKSPALSSWGDVPVLLKTDVPLSSAEEA
ECHWADTELNRNRRRRFCCKVEGYGSVCCKDPTPIEFSPDPLPDNKVLNVPVAVIAGNRPNYL
YRMLRSLLSAQGVSPQMITVFIDGYEPMDDVVALFGLRGIQHTPISIKNARVSQHYKASLTA
TFNLFPEAKFAVVLEEDLDIAVDFFSFLSQSIHLEEDDSLYCISAWNDAQYEHTAEDPALLY
RVETMPGLGWVLRRLYKEELEPKWPTPEKLWDWDMWMMRMEQRRGRECIIPDVSRSYHFGIV
GLNMNGYFHEAYFKKHKFNTVPGVQLRNVDLKEAYEVEVHRLLEAEVLDHKNPCEDSFL
PDTEGHTYVAFIRMEKDDDFTTWTQLAKCLHIWDLVVRGNHRGLWRLFRKKNHFLVVGVPASP
YSVKKPPSVTPIFLEPPPKEEGAPGAPEQT

Important features:**Transmembrane domain:**

amino acids 38-55

Homologous region to Mouse GNT1

amino acids 229-660

481/550

FIGURE 481

GAAAGAATGTTGTGGCTGCTCTTTTTCTGGTGACTGCCATTCATGCTGAACTCTGTCAACCA
GGTGCAGAAAATGCTTTTAAAGTGAGACTTAGTATCAGAACAGCTCTGGGAGATAAAGCATAT
GCCTGGGATACCAATGAAGAATACCTCTTCAAAGCGATGGTAGCTTTCTCCATGAGAAAAGTT
CCCAACAGAGAAGCAACAGAAATTTCCCATGTCCTACTTTGCAATGTAACCCAGAGGGTATCA
TTCTGGTTTGTGGTTACAGACCCTTCAAAAAATCACACCCTTCCTGCTGTTGAGGTGCAATCA
GCCATAAGAATGAACAAGAACCGGATCAACAATGCCTTCTTTCTAAATGACCAAACCTCTGGAA
TTTTTAAAAATCCCTTCCACACTTGCACCACCCATGGACCCATCTGTGCCCATCTGGATTATT
ATATTTGGTGTGATATTTTGCATCATCATAGTTGCAATTGCACTACTGATTTTATCAGGGATC
TGGCAACGTAGAAGAAAGAACAAAGAACCATCTGAAGTGGATGACGCTGAAGATAAGTGTGAA
AACATGATCACAATTGAAAATGGCATCCCCCTGATCCCCCTGGACATGAAGGGGGGCATATTA
ATGATGCCTTCATGACAGAGGATGAGAGGCTCACCCCTCTCTGAAGGGCTGTTGTTCTGCTTC
CTCAAGAAATTAAACATTTGTTTCTGTGTGACTGCTGAGCATCCTGAAATACCAAGAGCAGAT
CATATATTTTGTTCACCATTCCTCTTTTGTAAATAAATTTTGAATGTGCTTGAAAGTGAAAAG
CAATCAATTATACCCACCAACACCACTGAAATCATAAGCTATTCACGACTCAAATATTCTAA
AATATTTTCTGACAGTATAGTGTATAAATGTGGTCATGTGGTATTTGTAGTTATTGATTTAA
GCATTTTGTAGAAATAAGATCAGGCATATGTATATATTTTCACACTTCAAAGACCTAAGGAAAA
ATAAATTTTCCAGTGGAGAATACATATAATATGGTGTAGAAATCATTGAAAATGGATCCTTTT
TGACGATCACTTATATCACTCTGTATATGACTAAGTAAACAAAAGTGAGAAGTAATTATTGTA
AATGGATGGATAAAAATGGAATTACTCATATACAGGTGGAATTTTATCCTGTTATCACACCA
ACAGTTGATTATATATTTTCTGAATATCAGCCCCTAATAGGACAATTCTATTTGTTGACCATT
TCTACAATTTGTAAAAGTCCAATCTGTGCTAACTTAATAAAGTAATAATCATCTCTTTTAA
AAAAAAAAAAAAAAAAAAAAA

482/550

FIGURE 482

MLWLLFFLVTAIHAELCQPGAENAFKVRLSIRTALGDKAYAWDTNEEYLFKAMVAFSMRKVPN
REATEISHVLLCNVTQRVSEFWFVVTDP SKNHTLPAVEVQSAIRMNKNRINNAFFLNDQTLEFL
KIPSTLAPPMDPSVPIWIIIFGVIFCIIIVAIALLLSGIWQRRRKNKEPSEVDDAEDKCENM
ITIENGIPSDPLDMKGGILMMPS

483/550

FIGURE 483

CGTCTCTGCGTTTCGCC**ATG**CGTCCCGGGGCGCCAGGGCCACTCTGGCCTCTGCCCTGGGGGGC
CCTGGCTTGGGCCGTGGGCTTCGTGAGCTCCATGGGCTCGGGGAACCCCGCGCCCGGTGGTGT
TTGCTGGCTCCAGCAGGGCCAGGAGGCCACCTGCAGCCTGGTGCTCCAGACTGATGTACCCCG
GGCCGAGTGCTGTGCCTCCGGCAACATTGACACCGCCTGGTCCAACCTCACCCACCCGGGGAA
CAAGATCAACCTCCTCGGCTTCTTGGGCCTTGTCCACTGCCTTCCCTGCAAAGATTTCGTGCGA
CGGCGTGGAGTGCGGGCCCGGGCAAGGCGTGCCGCATGCTGGGGGGCCGCCCGCGCTGCGAGTG
CGCGCCCGACTGCTCGGGGCTCCCGGCGCGGCTGCAGGTCTGCGGCTCAGACGGCGCCACCTA
CCGCGACGAGTGCGAGCTGCGCGCCGCGCGCTGCCGCGGCCACCCGGACCTGAGCGTCATGTA
CCGGGGCCCGCTGCCGCAAGTCCTGTGAGCACGTGGTGTGCCCGCGGCCACAGTCGTGCGTTCGT
GGACCAGACGGGCAGCGCCCACTGCGTGGTGTGTGAGCGGGCGCCCTGCCCTGTGCCCTCCAG
CCCCGGCCAGGAGCTTTGCGGCAACAACAACGTCACCTACATCTCCTCGTGCCACATGCGCCA
GGCCACCTGCTTCCTGGGCCGCTCCATCGGCGTGCGCCACGCGGGCAGCTGCGCAGGCACCCC
TGAGGAGCCGCCAGGTGGTGAGTCTGCAGAAGAGGAAGAGAACTTCGTG**TGA**GCCTGCAGGAC
AGGCCTGGGCCTGGTGCCCCGAGGCCCCCCATCATCCCCTGTTATTTATTGCCACAGCAGAGTC
TAATTTATATGCCACGGACACTCCTTAGAGCCCGGATTTCGGACCACTTGGGGATCCCAGAAC
TCCCTGACGATATCCTGGAAGGACTGAGGAAGGGAGGCCTGGGGGCCGGCTGGTGGGTGGGAT
AGACCTGCGTTCCGGACACTGAGCGCCTGATTTAGGGCCCTTCTCTAGGATGCCCCAGCCCCCT
ACCCTAAGACCTATTGCCGGGGAGGATTCCACACTTCCGCTCCTTTGGGGATAAACCTATTAA
TTATTCTACTATCAAGAGGGCTGGGCATTCTCTGCTGGTAATTCCTGAAGAGGCATGACTGC
TTTTCTCAGCCCCAAGCCTCTAGTCTGGGTGTGTACGGAGGGTCTAGCCTGGGTGTGTACGGA
GGGTCTAGCCTGGGTGAGTACGGAGGGTCTAGCCTGGGTGAGTACGGAGGGTCTAGCCTGGGT
GAGTACGGAGGGTCTAGCCTGGGTGTGTATGGAGGATCTAGCCTGGGTGAGTATGGAGGGTCT
AGCCTGGGTGAGTATGGAGGGTCTAGCCTGGGTGTGTATGGAGGGTCTAGCCTGGGTGAGTAT
GGAGGGTCTAGCCTGGGTGTGTATGGAGGGTCTAGCCTGGGTGAGTATGGAGGGTCTAGCCTG
GGTGTGTACGGAGGGTCTAGTCTGAGTGCGTGTGGGGACCTCAGAACACTGTGACCTTAGCCC
AGCAAGCCAGGCCCTTCATGAAGGCCAAGAAGGCTGCCACCATTCCCTGCCAGCCCAAGAACT
CCAGCTTCCCCACTGCCTCTGTGTGCCCTTTGCGTCCTGTGAAGGCCATTGAGAAATGCCCA
GTGTGCCCCCTGGGAAAGGGCACGGCCTGTGCTCCTGACACGGGCTGTGCTTGGCCACAGAAC
CACCCAGCGTCTCCCCTGCTGCTGTCCACGTCAGTTCATGAGGCAACGTCGCGTGGTCTCAGA
CGTGAGCAGCCAGCGGCAGCTCAGAGCAGGGCACTGTGTCCGGCGGAGCCAAGTCCACTCTG
GGGAGCTCTGGCGGGGACCACGGGCCACTGCTCACCCACTGGCCCCGAGGGGGGTGTAGACG
CCAAGACTCACGCATGTGTGACATCCGGAGTCTGGAGCCGGGTGTCCAGTGGCACCCTAG
GTGCTGCTGCCTCCACAGTGGGGTTACACCCAGGGCTCCTTGGTCCCCCACAACCTGCCCC
GGCCAGGCCTGCAGACCCAGACTCCAGCCAGACCTGCCTCACCCACCAATGCAGCCGGGGCTG
GCGACACCAGCCAGGTGCTGGTCTTGGGCCAGTTCTCCACGACGGCTCACCTCCCTCCAT
CTGCGTTGATGCTCAGAATCGCCTACCTGTGCCTGCGTGTAAACCACAGCCTCAGACCAGCTA
TGGGGAGAGGACAACACGGAGGATATCCAGCTTCCCCGGTCTGGGGTGAGGAATGTGGGGAGC
TTGGGCATCCTCCTCCAGCCTCCTCCAGCCCCAGGCAGTGCCTTACCTGTGGTGCCAGAAA
AGTGCCCTAGGTGGTGGGTCTACAGGAGCCTCAGCCAGGCAGCCACCCACCCCTGGGGCC
CTGCTCACCAAGGAAATAAAGACTCAAGCCATAAAAAAA

484/550

FIGURE 484

MRPGAPGPLWPLPWGALAWAVGFVSSMSGSGNPAPGGVCWLQQGQEATCSLVLQTDVTRAEC
CA
SGNIDTAWSNLTHPGNKINLLGFLGLVHCLPCKDSCDGVCEGPGKACRMLGGRPRCECAPDCS
GLPARLQVCGSDGATYRDECELRAARCRGHPDLSVMYRGRCRKSCEHVVCPRPQSCVVDQTGS
AHCVVCRAAPCPVPSSPGQELCGNNNVTYISSCHMRQATCFLGRSIGVRHAGSCAGTPEEPPG
GESAEEEENFV

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation sites.

amino acids 73-77, 215-219

Osteonectin domain proteins.

amino acids 97-130, 169-202

485/550

FIGURE 485

GCTCGAGGCCGGCGGGCGGGGAGAGCGACCCGGGCGGCCTCGTAGCGGGGCCCCGGATCCCC
GAGTGGCGGCCGGAGCCTCGAAAAGAGATTCTCAGCGCTGATTTTGAGATGATGGGCTTGGA
AACGGGCGTCGCAGCATGAAGTCGCCGCCCTCGTGCTGGCCGCCCTGGTGGCTGCATCATC
GTCTTGGGCTTCAACTACTGGATTGCGAGCTCCCGGAGCGTGGACCTCCAGACACGGATCATG
GAGCTGGAAGGCAGGGTCCGCAGGGCGGCTGCAGAGAGAGGCGCCGTGGAGCTGAAGAAGAAC
GAGTTCCAGGGAGAGCTGGAGAAGCAGCGGGAGCAGCTTGACAAAATCCAGTCCAGCCACAAC
TTCCAGCTGGAGAGCGTCAACAAGCTGTACCAGGACGAAAAGGCGGTTTTGGTGAATAACATC
ACCACAGGTGAGAGGCTCATCCGAGTGCTGCAAGACCAGTTAAAGACCCTGCAGAGGAATTAC
GGCAGGCTGCAGCAGGATGTCCTCCAGTTTCAGAAGAACCAGACCAACCTGGAGAGGAAGTTC
TCCTACGACCTGAGCCAGTGCATCAATCAGATGAAGGAGGTGAAGGAACAGTGTGAGGAGCGA
ATAGAAGAGGTACCAAAAAGGGGAATGAAGCTGTAGCTTCCAGAGACCTGAGTGAAAACAAC
GACCAGAGACAGCAGCTCCAAGCCCTCAGTGAGCCTCAGCCCAGGCTGCAGGCAGCAGGCCTG
CCACACACAGAGGTGCCACAAGGGAAGGGAACGTGCTTGGTAACAGCAAGTCCCAGACACCA
GCCCCCAGTTCCGAAGTGGTTTTGGATTCAAAGAGACAAGTTGAGAAAGAGGAAACCAATGAG
ATCCAGGTGGTGAATGAGGAGCCTCAGAGGGACAGGCTGCCGCAGGAGCCAGGCCGGGAGCAG
GTGGTGGAAGACAGACCTGTAGGTGGAAGAGGCTTCGGGGGAGCCGGAGAACTGGGGCCAGACC
CCACAGGTGCAGGCTGCCCTGTCAGTGAGCCAGGAAAATCCAGAGATGGAGGGCCCTGAGCGA
GACCAGCTTGTCATCCCCGACGGACAGGAGGAGGAGCAGGAAGCTGCCGGGGAAGGGAGAAAC
CAGCAGAACTGAGAGGAGAAGATGACTACAACATGGATGAAAATGAAGCAGAATCTGAGACA
GACAAGCAAGCAGCCCTGGCAGGGAATGACAGAAACATAGATGTTTTAATGTTGAAGATCAG
AAAAGAGACACCATAAATTTACTTGATCAGCGTGAAAAGCGGAATCATACACTCTGAATTGAA
CTGGAATCACATATTTACAACAGGGCCGAAGAGATGACTATAAAATGTTTCATGAGGGACTGA
ATACTGAAAACCTGTGAAATGTACTAAATAAAATGTACATCTGA

486/550

FIGURE 486

MMGLGNRRSMKSPPLVLAALVACIIIVLGFNYWIIASSRSVDLQTRIMELEGRVRRAAAERGAV
ELKKNEFQGELEKQREQLDKIQSSHNFQLESVNKLYQDEKAVLVNNITTGERLIRVLQDQLKT
LQRYGRLQQDVLQFQKNQTNLERKFSYDLSQCINQMKEVKEQCEERIEEVTKKGNEAVASRD
LSENNDQRQQQLQALSEPQPRQLQAAGLPHTTEVPQGKGNVLGNSKSQTPAPSSEVVLDSCRQVEK
EETNEIQVVNEEPQRDRLPQEPGREQVVEDRPVGGRGFGGAGELGQTPQVQAALSVSQENPEM
EGPERDQLVIPDGQEEEQEAAGEGRNQKLRGEDDYNMDENEAESSETDKQAALAGNDRNIDVF
NVEDQKRDTINLLDQREKRNHTL

Important features:**Signal peptide:**

amino acids 1-29

487/550

FIGURE 487

AACTCAAACCTCTCTCTCTGGGAAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGGG
TGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTC**ATG**TAT
GGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATAACAGCTCACAGCTCTTTGGCCT
ATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCTGTTAATGGGACAGATGCTCGG
TTAAAATGCACTTTCTCCAGCTTTGCCCCGTGTGGGTGATGCTCTAACAGTGACCTGGAATTTT
CGTCCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTCCAACCC
ATGAGTGGGCGGTTTAAGGACCGGGTGTCTTGGGATGGGAATCCTGAGCGGTACGATGCCTCC
ATCCTTCTCTGGAAACTGCAGTTCGACGACAATGGGACATACACCTGCCAGGTGAAGAACCCA
CCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTCAGCGTCGTGCACACTGTACGCTTCTCT
GAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGATCATAATAGTAATT
GTAGTGGTCCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTCATAAAGTGGTG
GAGATAAAATCAAAGAAGAGGAAAGGCTCAACCAAGAGAAAAAGGTCTCTGTTTATTTAGAA
GACACAGAC**TAA**CAATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAAGAACCCTAGTATT
TCTTGAAGTTAATGGAACTTTTCTTTGGCTTTTCCAGTTGTGACCCGTTTCCAACCAGTTC
TGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCACAGTGCTCCTCCATAT
CACCAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTTCAGAGTGTAATTTTTTCAA
GTGCTCATTAGGTTTTATAAACAAGAAGCTACATTTTTGCCCTTAAGACACTACTTACAGTGT
TATGACTTGTATACACATATATTGGTATCAAAGGGGATAAAAGCCAATTTGTCTGTTACATTT
CCTTTCACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTAATGTGTTTACTCTCTTTCC
TTCCACATTCTCAATTAAAAGGTGAGCTAAGCCTCCTCGGTGTTTCTGATTAAACAGTAAATC
CTAAATTCAAACCTGTTAAATGACATTTTTATTTTTATGTCTCTCCTTAACATATGAGACACATC
TTGTTTTACTGAATTTCTTTCAATATTCCAGGTGATAGATTTTTGTCTG

488/550

FIGURE 488

MYGKSSTRAVLLLLGIQLTALWPAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTVTW
NFRPLDGGPEQFVFYYHIDPFQPMGRFKDRVSWDGNPERYDASILLWKLQFDDNGTYTCQVK
NPPDVGVGIVEIRLSVVHTVRFSEIHFLALAIGSACALMIIIVIVVVLQHYRKKRWAERAHK
VVEIKSKEEERLNQEKKVSVYLETD

FIGURE 489

[illegible]

490/550

FIGURE 490

MLLLWVSVVAALALAVLAPGAGEQRRRAAKAPNVVLVVSDSFDGRLTFHPGSQVVKLPFFINFM
KTRGTSFLNAYTNSPICCPSRAAMWSGLFTHLTESWNNFKGLDPNYTTWMDVMERHGYRTQKF
GKLDYTSBGHHSISNRVEAWTRDVAFLLRQEGRPMVNLIRNRTKVRVMERDWQNTDKAVNWLK
EAINYTEPFVIYLGLNLPHYPSPSSGENFGSSTFHTSLYWLEKVSHDAIKIPKWSPLSEMHP
VDYYSSYTKNCTGRFTKKEIKNIRAFYYAMCAETDAMLGEIILALHQLDLLQKTIVIIYSSDHG
ELAMEHRQFYKMSMYEASAHVPLLMMGPGIKAGLQVSNVSLVDIYPTMLDIAGIPLPQNLSG
YSLPLSSETFKNEHKVKNLHPPWILSEFHGCNVNASTYMLRTNHWKYIAYS DGASILPQLFD
LSSDPDELTVAVKFPEITYSLDQKLHSIINYPKVSASVHQYNKEQFIKWKQSIGQNYSNVIA
NLRWHQDWQKEPRKYENAIQWLKTHMNPRAV

Important features:**Signal peptide:**

amino acids 1-15

N-glycosylation sites.amino acids 108-111, 166-169, 193-196, 262-265, 375-378, 413-416,
498-501**Sulfatases proteins:**

amino acids 286-315, 359-369, 78-97

491/550

FIGURE 491

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGA
GCTTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGC
CTCTCTTGGCCTCCAACCTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGT
TGCCATGCTGCTCCCCAGCTGGAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGT
TGGCTTCTCCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGA
CATCTATAGCACCTTCTGGGCCTGCCCCTGACATCCAGGCTGCCAGGCCATGATGGTGAC
ATCCAGTGCAATCTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTT
CTGCCAGGAATCCCGAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTCATCCTTGG
AGGCCTCCTGGGATTCAATTCCTGTTGCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTC
ACCACTGGTGCCTGACAGCATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTT
TTCCCTGTTCTCCCTGATAGCTGGAATCATCCTCTGCTTTTCCTGCTCATCCCAGAGAAATCG
CTCCAATACTACTACGATGCCTACCAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGG
TCAACCTCCCAAAGTCAAGAGTGAGTTCAATTCCTACAGCCTGACAGGGTATGTGTGAAAGAAC
CAGGGGCCAGAGCTGGGGGGTGGCTGGGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCA
CAGGTGAGGGACACTACCACTGGATCGTGTGAGAAGGTGCTGCTGAGGATAGACTGACTTTGG
CCATTGGATTGAGCAAAGGCAGAAATGGGGGCTAGTGTAACAGCATGCAGGTTGAATTGCCAA
GGATGCTCGCCATGCCAGCCTTTCTGTTTTCTCACCTTGCTGCTCCCCTGCCCTAAGTCCCC
AACCCTCAACTTGAAACCCCATTCCTTAAGCCAGGACTCAGAGGATCCCTTTGCCCTCTGGT
TTACCTGGGACTCCATCCCCAAACCCACTAATCACATCCCACTGACTGACCCCTCTGTGATCAA
AGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATTGCTGGGGATGGGAAGGAGAAGCAGT
GGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTCCCTCCAAAGAACTGATTGGCCC
TGGAACCTCCATCCCACTCTTGTTATGACTCCACAGTGTCCAGACTAATTTGTGCATGAACTG
AAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATGCAGGATGGGAGGACAGGAA
GGCAGCCTGGGACATTTAAAAAATA

492/550

FIGURE 492

MASLGLQLVG YILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQ
CDIYSTLLGLPADIQAAQAMMVTSSAIISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFI
LGLLGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQR
NRSNYYDAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

Important features:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

493/550

FIGURE 493

GCACTGCTGCTGTCCCATCAGCTGCTCTGAAGCTCCATGGTGCCCAGAATCTTCGCTCCTGCT
TATGTGTCAGTCTGTCTCCTCCTCTTGTGTCCAAGGGAAGTCATCGCTCCCGCTGGCTCAGAA
CCATGGCTGTGCCAGCCGGCACCCAGGTGTGGAGACAAGATCTACAACCCCTTGGAGCAGTGC
TGTTACAATGACGCCATCGTGTCCCTGAGCGAGACCCGCCAATGTGGTCCCCCTGCACCTTC
TGGCCCTGCTTTGAGCTCTGCTGTCTTGATTCCCTTTGGCCTCACAAACGATTTTGTGTGAAG
CTGAAGGTTTCAGGGTGTGAATTCAGTGCCACTCATCTCCCATCTCCAGTAAATGTGAAAGC
AGAAGACGTTTTCCCTTGAGAAGACATAGAAAGAAAATCAACTTTCCTAAGGCATCTCAGAAA
CATAGGCTAAGGTAATATGTGTACCAGTAGAGAAGCCTGAGGAATTTACAAAATGATGCAGCT
CCAAGCCATTGTATGGCCCATGTGGGAGACTGATGGGACATGGAGAATGACAGTAGATTATCA
GGAAATAAATAAAGTGGTTTTTCCAATGTACACACCTGTAAAA

494/550

FIGURE 494

MVPRI FAPAYVSVCLLLCPREVIAPAGSEPWLCQPAPRCGDKIYNPLEQCCYNDAIVSLSET
RQCGPPCTFWPCFELCCLDSFGLTNDFFVVKLKVQGVNSQCHSSPISSKCESRRRFP

Important features:

Signal peptide:

amino acids 1-25

495/550

FIGURE 495

CTCCACTGCAACCACCCAGAGCCATGGCTCCCCGAGGCTGCATCGTAGCTGTCTTTGCCATTT
TCTGCATCTCCAGGCTCCTCTGCTCACACGGAGCCCCAGTGGCCCCCATGACTCCTTACCTGA
TGCTGTGCCAGCCACACAAGAGATGTGGGGACAAGTTCTACGACCCCCTGCAGCACTGTTGCT
ATGATGATGCCGTCGTGCCCTTGCCAGGACCCAGACGTGTGGAACTGCACCTTCAGAGTCT
GCTTTGAGCAGTGCTGCCCCTGGACCTTCATGGTGAAGCTGATAAACCAGAACTGCGACTCAG
CCCGGACCTCGGATGACAGGCTTTGTGCGAGTGTGAGCTAATGGAACATCAGGGGAACGATGA
CTCCTGGATTCTCCTTCCTGGGTGGGCCTGGAGAAAGAGGCTGGTGTTACCTGAGATCTGGGA
TGCTGAGTGGCTGTTTGGGGGCCAGAGAAACACACACTCAACTGCCCACTTCATTCTGTGACC
TGTCTGAGGCCCACCCTGCAGCTGCCCTGAGGAGGCCCACAGGTCCCCTTCTAGAATTCTGGA
CAGCATGAGATGCGTGTGCTGATGGGGGCCAGGGACTCTGAACCCTCCTGATGACCCCTATG
GCCAACATCAACCCGGCACCACCCCAAGGCTGGCTGGGGAACCCTTCACCCTTCTGTGAGATT
TTCCATCATCTCAAGTTCTTCTATCCAGGAGCAAAGCACAGGATCATAATAAATTTATGTA
CTTTATAAATGAAAA

496/550

FIGURE 496

MAPRGCIVAVFAIFCISRLLC SHGAPVAPMTPYLMLCQPHKRCGDKFYDPLQHCCYDDAVVPL
ARTQTCGNCTFRVCFEQCCPWTFMVKLINQNCDSARTSDDRLLCRSVS

Important features:

Signal peptide:

amino acids 1-24

FIGURE 497

TGAAGGACTTTTCCAGGACCCAAGGCCACACACTGGAAGTCTTGACGCTGAAGGGAGGCACCT
CTTGGCCTCCGCAGCCGATCAC**ATGA**AAGGTGGTGCCAAGTCTCCTGCTCTCCGTCTCTGGC
ACAGGTGTGGCTGGTACCCGGCTTGCCCCCAGTCTCAGTCGCCAGAGACCCCAGCCCCCA
GAACCAGACCAGCAGGGTAGTGCAGGCTCCCAGGGAGGAAGAGGAAGATGAGCAGGAGGCCAG
CGAGGAGAAGGCCGGTGAGGAAGAGAAAGCCTGGCTGATGGCCAGCAGGCAGCAGCTTGCCAA
GGAGACTTCAAACCTTCGATTACGCCTGCTGCGAAAGATCTCCATGAGGCACGATGGCAACAT
GGTCTTCTCTCCATTTGGCATGTCTTGCCATGACAGGCTTGATGCTGGGGGCCACAGGGCC
GACTGAAACCCAGATCAAGAGAGGGGCTCCACTTGCCAGGCCCTGAAGCCCACCAAGCCCCGGGCT
CCTGCCTTCCCTCTTTAAGGGACTCAGAGAGACCCTCTCCCGCAACCTGGA**ACT**GGGCCTCTC
ACAGGGGAGTTTTGCCTTCATCCACAAGGATTTTGATGTCAAAGAGACTTTCTTCAATTTATC
CAAGAGGTATTTTGATACAGAGTGCCTGCTATGAATTTTCGCAATGCCCTCACAGGCCAAAAG
GCTCATGAATCATTACATTAACAAAGAGACTCGGGGGAAAATTTCCCAA**ACT**GTTTGATGAGAT
TAATCCTGAAACCAAATTAATTCTTGTTGGATTACATCTTGTTCAAAGGGAAATGGTTGACCCC
ATTTGACCCTGTCTTCACCGAAGTCGACACTTTCCACCTGGACAAGTACAAGACCATTAAGGT
GCCCCATGATGTACGGTGCAGGCAAGTTTGCCCTCCACCTTGACAAGAATTTTCGTTGTCATGT
CCTCAAACCTGCCCTACCAAGGAAATGCCACCATTGCTGGTGGTCTCATGGAGAAAATGGTGTA
CCACCTCGCCCTTGAAGACTACCTGACCACAGACTTGGTGGAGACATGGCTCAAGAAACATGA
AACCAGAAAACATGGAAGTTTTCTTTCCGAAGTTCAAGCTAGATCAGAAGTATGAGATGCATGA
GCTGCTTAGGCAGATGGGAATCAGAAGAATCTTCTCACCCCTTGCTGACCTTAGTGA**ACT**CTC
AGCTACTGGAAGAAATCTCCAAGTATCCAGGGTTTTACGAAGAACAGTGATTGAAGTTGATGA
AAGGGGCACTGAGGCAGTGGCAGGAATCTTGTCAGAAATTACTGCTTATTCCATGCCTCCTGT
CATCAAAGTGGAACCGCCATTTCAATTCATGATCTATGAAGAAACCTCTGGAATGCTTCTGTT
TCTGGGCAGGGTGGTGAATCCGACTCTCCTA**TA**ATTTCAGGACATGCATAAGCACTTCGTGCTG
TAGTAGATGCTGAATCTGAGGTATCAAACACACACAGGATACCAGCAATGGATGGCAGGGGAG
AGTGTTCCCTTTTGTTCTTAAGTAGTTTAGGGTGTTCTCAAATAAATACAGTAGTCCCCACTTA
TCTGAGGGGGATACATTCAAAGACCCCCAGCAGATGCCTGAAACGGTGGACAGTGCTGAACCT
TATATATATTTTTTCTACACATACATACCTATGATAAAAGTTTAATTTATAAATTAGGCACAG
TAAGAGATTAACAATAATAACAACATTAAGTAAATGAGTTACTTGAACGCAAGCACTGCAAT
ACCATAACAGTCAAACCTGATTATAGAGAAGGCTACTAAGTGACTCATGGGCGAGGAGCATAGA
CAGTGTGGAGACATTGGGCAAGGGGAGAATTCACATCCTGGGTGGGACAGAGCAGGACGATGC
AAGATTCCATCCCCTACTCAGAATGGCATGCTGCTTAAGACTTTTAGATTGTTTATTTCTG
AATTTTTCTATTAAATGTTTTTGGACCATGGTTGACCATGGTTAACTGAGACTGCAGAAAGCAA
AACCATTGGATAAGGGAGGACTACTACAAAAGCATTAATTTGATACATATTTTTTAAAA**AAAA**
AAAA**AAAA**

498/550

FIGURE 498

MKVVP S L L L S V L L A Q V W L V P G L A P S P Q S P E T P A P Q N Q T S R V V Q A P R E E E E D E Q E A S E E K A G E E
E K A W L M A S R Q Q L A K E T S N F G F S L L R K I S M R H D G N M V F S P F G M S L A M T G L M L G A T G P T E T Q I K R
G L H L Q A L K P T K P G L L P S L F K G L R E T L S R N L E L G L S Q G S F A F I H K D F D V K E T F F N L S K R Y F D T E
C V P M N F R N A S Q A K R L M N H Y I N K E T R G K I P K L F D E I N P E T K L I L V D Y I L F K G K W L T P F D P V F T E
V D T F H L D K Y K T I K V P M M Y G A G K F A S T F D K N F R C H V L K L P Y Q G N A T M L V V L M E K M G D H L A L E D Y
L T T D L V E T W L R N M K T R N M E V F F P K F K L D Q K Y E M H E L L R Q M G I R R I F S P F A D L S E L S A T G R N L Q
V S R V L R R T V I E V D E R G T E A V A G I L S E I T A Y S M P P V I K V D R P F H F M I Y E E T S G M L L F L G R V V N P
T L L

FIGURE 499

[illegible]

500/550

FIGURE 500

MDSLRLKMLISVAMLGAGAGVGYALLVIVTPGERRKQEMLKEMPLQDPRSREEAARTQQLLLAT
LQEAATTQENVAWRKNWMVGEGGASGRSP

Important features:

Signal peptide:

amino acids 1-18

501/550

FIGURE 501

[illegible]

502/550

FIGURE 502

MGPSTPLLILFLLSWGPLQGQHHLVEYMERRLAALERLAQCQDQSSRHAAELRDFKNKML
PLLEVAEKEREALRTEADTISGRVDRLEREVDYLETQNPALPCVEFDEKVTGGPGTKGKGRN
EKYDMVTDCGYTISQVRSMKILKREGGPAGLWTKDPLGQTEKIYVLDGTQNDTAFVFPRLRDF
TLAMAARKASRVRVFPWPVGTGQLVYGGFLYFARRPPGRPGGGGEMENTLQLIKFHLANRTVV
DSSVFPAEGLIPPYGLTADTYIDLVADEEGLWAVYATREDDRHLCCLAKLDPQTLDTQQWDTP
CPRENAEAAFVICGTLYVVYNTRPASRARIQCSFDASGTLTPERAALPYFPRRYGAHASLRYN
PRERQLYAWDDGYQIVYKLEMRKKEEV

Important features:**Signal peptide:**

amino acids 1-21

N-glycosylation sites.

amino acids 177-180, 248-251

503/550

FIGURE 503

TGCGGCGCAGTGTAGACCTGGGAGG**ATG**GGCGGCCTGCTGCTGGCTGCTTTTCTGGCTTTGGT
CTCGGTGCCCAGGGCCCAGGCCGTGTGGTTGGGAAGACTGGACCCTGAGCAGCTTCTTGGGCC
CTGGTACGTGCTTGCGGTGGCCTCCCGGGAAAAGGGCTTTGCCATGGAGAAGGACATGAAGAA
CGTCGTGGGGGTGGTGGTGACCCTCACTCCAGAAAACAACCTGCGGACGCTGTCCTCTCAGCA
CGGGCTGGGAGGGTGTGACCAGAGTGTGATGGACCTGATAAAGCGAACTCCGGATGGGTGTT
TGAGAATCCCTCAATAGGCGTGCTGGAGCTCTGGGTGCTGGCCACCAACTTCAGAGACTATGC
CATCATCTTCACTCAGCTGGAGTTCGGGGACGAGCCCTTCAACACCGTGGAGCTGTACAGTCT
GACGGAGACAGCCAGCCAGGAGGCCATGGGGCTCTTACCAAGTGGAGCAGGAGCCTGGGCTT
CCTGTCACAG**TAG**CAGGCCCAGCTGCAGAAGGACCTCACCTGTGCTCACAAGATCCTTCTGTG
AGTGCTGCGTCCCCAGTAGGGATGGCGCCCACAGGGTCCTGTGACCTCGGCCAGTGTCCACCC
ACCTCGCTCAGCGGCTCCCGGGGCCAGCACCAGCTCAGAATAAAGCGATTCCACAGCA

504/550

FIGURE 504

·MGGLLLAFLALVSVPRQAVWLGRLDPEQLLGPWYVLAVASREKGFAMEKDMKNVVGVVTL
TPENNLRTLSSQHGLGGCDQSVMDLIKRN¹SGWVFENPSIGVLELWVLATNFRDYAIIFTQLEF
GDEPFNTVELYSLTETASQEAMGLFTKWSRSLGFLSQ

Important features:

Signal peptide:

amino acids 1-20

505/550

FIGURE 505

GTTCCGCAGATGCAGAGGTTGAGGTGGCTGCGGGACTGGAAGTCATCGGGCAGAGGTCTCACA
GCAGCCAAGGAACCTGGGGCCCGCTCCTCCCCCTCCAGGCCATGAGGATTCTGCAGTTAATC
CTGCTTGCTCTGGCAACAGGGCTTGTAGGGGGAGAGACCAGGATCATCAAGGGGTTCGAGTGC
AAGCCTCACTCCCAGCCCTGGCAGGCAGCCCTGTTGAGAAGACGCGGCTACTCTGTGGGGCG
ACGCTCATCGCCCCAGATGGCTCCTGACAGCAGCCCACTGCCTCAAGCCCCGCTACATAGTT
CACCTGGGGCAGCACAACTCCAGAAGGAGGAGGGCTGTGAGCAGACCCGGACAGCCACTGAG
TCCTTCCCCCACCCGGCTTCAACAACAGCCTCCCCAACAAAGACCACCGCAATGACATCATG
CTGGTGAAGATGGCATCGCCAGTCTCCATCACCTGGGCTGTGCGACCCCTCACCTCTCCTCA
CGCTGTGTCACTGCTGGCACCAGCTGCCTCATTTCCGGCTGGGGCAGCACGTCCAGCCCCCAG
TTACGCCTGCCTCACACCTTGCATGCGCCAACATCACCATCATTGAGCACCAGAAGTGTGAG
AACGCCTACCCCGGAACATCACAGACACCATGGTGTGTGCCAGCGTGCAGGAAGGGGGCAAG
GACTCCTGCCAGGGTGACTCCGGGGGCCCTCTGGTCTGTAACCAGTCTCTTCAAGGCATTATC
TCCTGGGGCCAGGATCCGTGTGCGATCACCCGAAAGCCTGGTGTCTACACGAAAGTCTGCAAA
TATGTGGACTGGATCCAGGAGACGATGAAGAACAATTAGACTGGACCCACCCACCACAGCCCA
TCACCCTCCATTTCCACTTGGTGTGTTGGTTCCTGTTCACTCTGTTAATAAGAAACCCTAAGCC
AAGACCCTCTACGAACATTCTTTGGGCCTCCTGGACTACAGGAGATGCTGTCACTTAATAATC
AACCTGGGGTTCGAAATCAGTGAGACCTGGATTCAAATTCTGCCTTGAAATATTGTGACTCTG
GGAATGACAACACCTGGTTTGTCTCTGTTGTATCCCCAGCCCCAAAGACAGCTCCTGGCCAT
ATATCAAGGTTTCAATAAATATTTGCTAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAA

506/550

FIGURE 506

MRILQLILLALATGLVGGETRIIKGFECKPHSQPWQAALFEKTRLLCGATLIAPRWLLTAAHC
LKPRYIVHLGQHNLQKEEGCEQTRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPV SITWAV
RPLTLSSRCV TAGTSC LISGWGSTSSPQLRLPHTLR CANIT IIEHQKCENAYPGNITDTMVCA
SVQEGGKDSCQGD SGGPLVCNQSLQGIISWGDPCA ITRKPGVYTKVCKYVDWIQETMKNN

Important features:**Signal peptide:**

amino acids 1-18

Serine proteases, trypsin family, histidine active site.

amino acids 58-63

N-glycosylation sites.

amino acids 99-102, 165-168, 181-184, 210-213

Glycosaminoglycan attachment site.

amino acids 145-148

Kringle domain proteins.

amino acids 197-209, 47-64

Serine proteases, trypsin family, histidine protein

amino acids 199-209, 47-63, 220-243

Apple domain proteins

amino acids 222-249, 189-222

507/550

FIGURE 507

CTGGGATCAGCCACTGCAGCTCCCTGAGCACTCTCTACAGAGACGCGGACCCCAGACATGAGG
AGGCTCCTCCTGGTCACCAGCCTGGTGGTTGTGCTGCTGTGGGAGGCAGGTGCAGTCCCAGCA
CCCAAGGTCCCTATCAAGATGCAAGTCAAACACTGGCCCTCAGAGCAGGACCCAGAGAAGGCC
TGGGGCGCCCGTGTGGTGGAGCCTCCGGAGAAGGACGACCAGCTGGTGGTGTGTTCCCTGTC
CAGAAGCCGAAACTCTTGACCACCGAGGAGAAGCCACGAGGTCAGGGCAGGGGCCCCATCCTT
CCAGGCACCAAGGCCTGGATGGAGACCGAGGACACCCTGGGGCCGTGTCCTGAGTCCCGAGCCC
GACCATGACAGCCTGTACCACCCTCCGCCTGAGGAGGACCAGGGCGAGGAGAGGCCCCGGTTG
TGGGTGATGCCAAATCACCAGGTGCTCCTGGGACCGGAGGAAGACCAAGACCACATCTACCAC
CCCCAGTAGGGCTCCAGGGGCCATCACTGCCCCCGCCCTGTCCCAAGGCCCAGGCTGTTGGGA
CTGGGACCCTCCCTACCCTGCCCCAGCTAGACAAATAAACCCCAGCAGGCAAAAAAAAAAAAAA
AAAAAA

508/550

FIGURE 508

MRRLLLVTSLVVVLLWEAGAVPAPKVPIKMQVKHWPSEQDPEKAWGARVVEPPEKDDQLVVLF
PVQKPKLLTTEEKPRGQGRGPILPGTKAWMETEDTLGRVLSPEPDHDSLYHPPPEEDQGEERP
RLWVMPNHQVLLGPEEDQDHIYHPQ

509/550

FIGURE 509

GCGGAGCCGGCGCCGGCTGCGCAGAGGAGCCGCTCTCGCCGCCGCCACCTCGGCTGGGAGCCC
ACGAGGCTGCCGCATCCTGCCCTCGGAACAATGGGACTCGGCGCGCGAGGTGCTTGGGCCGCG
CTGCTCCTGGGGACGCTGCAGGTGCTAGCGCTGCTGGGGGCCGCCCATGAAAGCGCAGCCATG
GCGGCATCTGCAAACATAGAGAATTCTGGGCTTCCACACAACCTCCAGTGCTAACTCAACAGAG
ACTCTCCAACATGTGCCTTCTGACCATACAAATGAAACTTCCAACAGTACTGTGAAACCACCA
ACTTCAGTTGCCTCAGACTCCAGTAATACAACGGTCACCACCATGAAACCTACAGCGGCATCT
AATACAACAACACCAGGGATGGTCTCAACAAATATGACTTCTACCACCTTAAAGTCTACACCC
AAAACAACAAGTGTTTCACAGAACACATCTCAGATATCAACATCCACAATGACCGTAACCCAC
AATAGTTCAGTGACATCTGCTGCTTCATCAGTAACAATCACAACAACCTATGCATTCTGAAGCA
AAGAAAGGATCAAAATTTGATACTGGGAGCTTTGTTGGTGGTATTGTATTAACGCTGGGAGTT
TTATCTATTCTTTACATTGGATGCAAAATGTATTACTCAAGAAGAGGCATTCGGTATCGAACC
ATAGATGAACATGATGCCATCATTTAAGGAAATCCATGGACCAAGGATGGAATACAGATTGAT
GCTGCCCTATCAATTAATTTTGGTTTATTAATAGTTTAAAACAATATTCTCTTTTTGAAAATA
GTATAAACAGGCCATGCATATAATGTACAGTGTATTACGTAAATATGTAAAGATTCTTCAAGG
TAACAAGGGTTTGGGTTTTGAAATAAACATCTGGATCTTATAGACCGTTCATACAATGGTTTT
AGCAAGTTCATAGTAAGACAAACAAGTCCTATCTTTTTTTTTTGGCTGGGGTGGGGGCATTGG
TCACATATGACCAGTAATTGAAAGACGTCATCACTGAAAGACAGAATGCCATCTGGGCATACA
AATAAGAAGTTTGTACAGCACTCAGGATTTTGGGTATCTTTTGTAGCTCACATAAAGAACTT
CAGTGCTTTTCAGAGCTGGATATATCTTAATTACTAATGCCACACAGAAATTATACAATCAAA
CTAGATCTGAAGCATAATTTAAGAAAAACATCAACATTTTTTGTGCTTTAACTGTAGTAGTT
GGTCTAGAAACAAAATACTCC

510/550

FIGURE 510

MGLGARGAWAALLLGLTLQVLALLGAAHESAAMAASANIENSGLPHNSSANSTETLQHVPDHT
NETSNSTVKPPTSVASDSSNTTVTTMKPTAASNTTTPGMVSTNMTSTTLKSTPKTTSVSQNTS
QISTSTMTVTHNSSVTSAASSVTITTTMHSEAKKGSKFDTGSEFVGGIVLTLGVLSILYIGCKM
YYSRRGIRYRTIDEHDAII

511/550

FIGURE 511

GACTTTGCTTGAATGTTTACATTTTCTGCTCGCTGTCCTACATATCACAATATAGTGTTACGTTTGTAAAC
TTTGGGGTGTGAGGAGTTGAGCTTGCTCAGCAAGCCAGCATGGCTAGGATGAGCTTTGTTATAGCAGCTTGCCAA
TTGGTGCTGGGCCTACTAATGACTTCATTAACCGAGTCTTCCATACAGAATAGTGAGTGCCACAACCTTTGCGTA
TGTGAAATTCGTCCCTGGTTTACCCACAGTCAACTTACAGAGAAGCCACCACTGTTGATTGCAATGACCTCCGC
TTAACAAGGATTCCCAGTAACCTCTCTAGTGACACACAAGTGCTTCTTACAGAGCAATAACATCGCGAAGACT
GTGGATGAGCTGCAGCAGCTTTTCAACTTGACTGAAGTAGATTTCTCCAAAACAACCTTTACTAACATTAAGGAG
GTCGGGCTGGCAAACCTAACCAGCTCACAACGCTGCATTGGAGGAAAATCAGATTACCGAGATGACTGATTAC
TGTCTACAAGACCTCAGCAACCTTCAAGAACTCTACATCAACCACAACCAAATTAGCACTATTTCTGCTCATGCT
TTTGCAGGCTTAAAAATCTATTAAGGCTCCACCTGAAGTCCAACTGAAAGTTATTGATAGTCGCTGGTTT
GATTCTACACCCAACCTGGAAATTCTCATGATCGGAGAAAACCTGTGATTGGAATTCTGGATATGAACCTTCAA
CCCCTCGCAAATTTGAGAAGCTTAGTTTTGGCAGGAATGTATCTCACTGATATTCCTGGAAATGCTTTGGTGGGT
CTGGATAGCCTTGAGAGCCTGTCTTTTATGATAACAACTGGTTAAAGTCCCTCAACTTGCCCTGCAAAAAGTT
CCAAATTTGAAATTCTTAGACCTCAACAAAAACCCATTCAACAAATCCAAGAAGGGGACTTCAAAAATATGCTT
CGGTTAAAGAAGTGGGAATCAACAATATGGGCGAGCTCGTTTCTGTGACCGCTATGCCCTGGATAACTTGCCT
GAAGTCAACAAAGCTGGAAGCCACCAATAACCCCTAACTCTCTTACATCCACCGCTTGGCTTTCGGAAGTGCCCT
GCTCTGGAAAGCTTGATGCTGAACAACAATGCCTTGAATGCCATTTACCAAAAGACAGTCGAATCCCTCCCCAAT
CTGCGTGAGATCAGTATCCATAGCAATCCCCTCAGGTGTGACTGTGTGATCCACTGGATTAAGTCCAAACAAACC
AACATCCGCTTCATGGAGCCCCTGTCCATGTTCTGTGCCATGCCGCCGAATATAAAGGGCACCAGGTGAAGGAA
GTTTTAATCCAGGATTCGAGTGAACAGTGCCTCCCAATGATATCTCACGACAGCTTCCCAATCGTTTAAACGTG
GATATCGGCACGACGGTTTTCTTAGACTGTGAGCCATGGCTGAGCCAGAACCTGAAATTTACTGGGTCACTCCC
ATTGGAAATAAGATAACTGTGGAAACCTTTAGATAAATACAAGCTAAGTAGCGAAGGTACCTTGAAATATCT
AACATACAAATTGAAGACTCAGGAAGATACACATGTGTTGCCAGAATGTCCAAGGGGCAGACACTCGGGTGGCA
ACAATTAAGGTTAACGGGACCTTCTGGATGGTACCCAGGTGCTAAAAATATACGTCAAGCAGACAGAATCCCAT
TCCATCTTAGTGTCTTGAAAGTTAATTCATGTGATGACGTCAAACTTAAATGGTCGTCTGCCACCATGAAG
ATTGATAACCTCACATAACATATACTGCCAGGGTCCCAGTGCATGTCCATGAATACAACCTAACGCATCTGCAG
CCTTCCACAGATTATGAAGTGTGTCTCACAGTGTCCAATATTCATCAGCAGACTCAAAAGTCATGCGTAAATGTC
ACAACCAAAATGCCGCCTTCGAGTGGACATCTCTGATCAAGAAACCAGTACAGCCCTTGCTGCAGTAATGGGG
TCTATGTTTGGCGTCATTAGCCTTGCGTCCATTGCTGTGTACTTTGCCAAAAGATTTAAGAGAAAAAACTACCAC
CACTCATTAAGAAAGTATATGCAAAAAACCTCTCAATCCCACTAAATGAGCTGTACCCACCACTCATTAACCTC
TGGGAAGGTGACAGCGAGAAAGACAAAGATGGTTCTGCAGACACCAAGCCAACCCAGGTGACACATCCAGAAGC
TATTACATGTGGTAACTCAGAGGATATTTGCTTCTGGTAGTAAGGAGCACAAAGACGTTTTTGTCTTTATCTGC
AAAAGTGAACAAGTTGAAGACTTTTGTATTTTGTACTTTGCTAGTTTGTGGCAGAGTGGAGAGGACGGGTGGATA
TTTCAAATTTTTTTAGTATAGCGTATCGCAAGGGTTTGACACGGCTGCCAGCGACTCTAGGCTTCCAGTCTGTGT
TTGGTTTTTATTCTTATCATTATATGATTGTTATTATATTATTTTATTTTAGTTGTTGTGCTAAACTCAAT
AATGCTGTCTAACTACAGTGCTCAATAAAATGATTAATGACAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

512/550

FIGURE 512

MARMSFVIAACQLVLGLLMTSLTESSIONSECPQLCVCEIRPWFTPQSTYREATTVDCNDLRL
TRIPSNLSSDTQVLLLQSNNIAKTVDELQQLFNLTEDFSQNNFTNIKEVGLANLTQLTTLHL
EENQITEMTDYCLQDLSNLQELYINHNQISTISAHAFAGLKNLLRLHLNSNKLKVIDSRWFDS
TPNLEILMIGENPVIGILDMNFKPLANLRSVLVLAGMYLTDIPGNALVGLDSLESLSFYDNKLV
KVPQLALQKVPNLKFLDLNKNPIHKIQEGDFKNMLRLKELGINNMGELVSVDRYALDNLPELT
KLEATNNPKLSYIHRLAFRSVPALLESMLNNAIYQKTVESLPNLREISIHNSNPLRCDV
IHWINSNKTNIRFMEPLSMFCAMPPEYKQVKEVLIQDSSEQCLPMISHDSFPNRLNVDIGT
TVFLDCRAMAEPEPEIYVWTPIGNKITVETLSDKYKLSSEGTLNISNIQIEDSGRYTCVAQNV
QGADTRVATIKVNGTLLDGTQVLKIYVKQTESHSILVSWKVNSNVMTSNLKWSSATMKIDNPH
ITYTARVPVDVHEYNLTHLQPDYEVCLTVSNIHQQTQKSCVNVTTKNAAFVAVDISDQETST
ALAAVMGSMFAVISLASIAVYFAKRFKRKNYHHSLLKKYMQKTSSIPLNELYPPPLINLWEGDSE
KDKDGSADTKPTQVDTSRSYMW

Important features:**Signal peptide:**

Amino acids 1-25

Transmembrane domain:

Amino acids 508-530

N-glycosylation sites:Amino acids 69-73;96-100;106-110;117-121;385-389;517-521;
582-586;611-615**Tyrosine kinase phosphorylation site:**

Amino acids 573-582

N-myristoylation sites:

Amino acids 16-22;224-230;464-470;637-643;698-704

513/550

FIGURE 513

GGGAGAGAGGATAAATAGCAGCGTGGCTTCCCTGGCTCCTCTCTGCATCCTTCCCGACCTTCC
CAGCAATATGCATCTTGCACGTCTGGTCGGCTCCTGCTCCCTCCTTCTGCTACTGGGGGCCCT
GTCTGGATGGGCGGCCAGCGATGACCCCATTGAGAAGGTCATTGAAGGGATCAACCGAGGGCT
GAGCAATGCAGAGAGAGAGGTGGGCAAGGCCCTGGATGGCATCAACAGTGGAATCACGCATGC
CGGAAGGGAAGTGGAGAAGGTTTTCAACGGACTTAGCAACATGGGGAGCCACACCGGCAAGGA
GTTGGACAAAGGCGTCCAGGGGCTCAACCACGGCATGGACAAGGTTGCCCATGAGATCAACCA
TGGTATTGGACAAGCAGGAAAGGAAGCAGAGAAGCTTGGCCATGGGGTCAACAACGCTGCTGG
ACAGGCCGGGAAGGAAGCAGACAAAGCGGTCCAAGGGTTCCACACTGGGGTCCACCAGGCTGG
GAAGGAAGCAGAGAACTTGGCCAAGGGGTCAACCATGCTGCTGACCAGGCTGGAAAGGAAGT
GGAGAAGCTTGGCCAAGGTGCCCACCATGCTGCTGGCCAGGCCGGAAGGAGCTGCAGAATGC
TCATAATGGGGTCAACCAAGCCAGCAAGGAGGCCAACCAGCTGCTGAATGGCAACCATCAAAG
CGGATCTTCCAGCCATCAAGGAGGGGCCACAACCACGCCGTTAGCCTCTGGGGCCTCAGTCAA
CACGCCTTTCATCAACCTTCCCGCCCTGTGGAGGAGCGTCGCCAACATCATGCCCTTAAACTGG
CATCCGGCCTTGCTGGGAGAATAATGTCGCCGTTGTCACATCAGCTGACATGACCTGGAGGGG
TTGGGGGTGGGGGACAGGTTTCTGAAATCCCTGAAGGGGGTGTACTGGGATTTGTGAATAAA
CTTGATACACCA

514/550

FIGURE 514

MHLARLVGSCSLLLLLGALSGWAASDDPIEKVIEGINRGLSNAEREVGKALDGINSGITHAGR
EVEKVFNGLSNMGSHGTGKELDKGVQGLNHGMDKVAHEINHGIGQAGKEAEKLGHGVNNAAGQA
GKEADKAVQGFHTGVHQAGKEAEKLGQGVNHAADQAGKEVEKLGQGAHHAAGQAGKELQNAHN
GVNQASKEANQLLNGNHQSGSSSHQGGATTTPLASGASVNTPPFINLPALWRSVANIMP

Important features:

Signal peptide:

amino acids 1-25

Homologous region to circumsporozoite (CS) repeats:

amino acids 35-225

515/550

FIGURE 515

CCCACGCGTCCGCCCACGCGTCCGGGTGCCACTCGCGCGCCGGCCGCGCTCCGGGCTTCTCTT
TTCCCTCCGACGCGCCACGGGTGCCAGACATTCCGGCTGCCGGGTCTGGAGAGCTCCCCGAA
CCCCTCCGCGGAGAGGAGCGAGGCGGCGCCAGGGTGGCCCCGGGGCGCGCTTGGTCTCGGAG
AAGCGGGGACGAGGCCGGAGGATGAGCGACTGAGGGCGACGCGGGCACTGACGCGAGTTGGGG
CCGCGACTACCGGCAGCTGACAGCGCGATGAGCGACTCCCCAGAGACGCCCTAGCCCGGTGTG
CGCGCCAGGCGGAGCGCGCAGGTGGGGCTGGGCTGTAGTGGTCCGCCCCACGCGGGTGC
GCCGGCCCAGGATGGGCGCTGGCAACCCGGGCGCGCCCGCCGCTGCTACCCCTGCGCCCGC
TGCGAGCCCGGCGTCCGGCCCCGCGCCCTGCGCTCATGGACGGCGGCTCCCGCTGGCGGCGGC
GCGCCCCCGGGCTGTGAATGCGACTCGCCCCCTCGGCCGCGCTCCCCGCCCCGCGCCGCGG
GACGTGGTAGGGGATGCCCCAGCTCCACTGCGATGGCAGTTGGCGCGCTCTCCAGTTCCCTCCT
GGTCACCTGCTGCCTGATGGTGGCTCTGTGCAGTCCGAGCATCCCGCTGGAGAAGCTGGCCCA
GGCACCAGAGCAGCCGGGCCAGGAGAAGCGTGAGCACGCCACTCGGGACGGCCCCGGGGCGGGT
GAACGAGCTCGGGCGCCCGGCGAGGGACGAGGGCGGACGCGGCCGGGACTGGAAGAGCAAGAG
CGGCCGTGGGCTCGCCGGCCGTGAGCCGTGGAGCAAGCTGAAGCAGGCCTGGGTCTCCCAGGG
CGGGGGCGCCAAGGCCGGGGATCTGCAGGTCCGGCCCCGCGGGGACACCCCGCAGGCGGAAGC
CCTGGCCGCAGCCGCCCAGGACGCGATTGGCCCGGAACTCGCGCCCACGCCCCAGCCACCCGA
GGAGTACGTGTACCCGGACTACCGTGGCAAGGGCTGCGTGGACGAGAGCGGCTTCGTGTACGC
GATCGGGGAGAAGTTCGCGCCGGGGCCCTCGGCCTGCCCGTGCCTGTGCACCGAGGAGGGGCC
GCTGTGCGCGCAGCCGAGTGCCCGAGGCTGCACCCGCGCTGCATCCACGTGCACACGAGCCA
GTGCTGCCCCGAGTGCAAGGAGAGGAAGAACTACTGCGAGTTCCGGGGCAAGACCTATCAGAC
TTTGGAGGAGTTCGTGGTGTCTCCATGCGAGAGGTGTCGCTGTGAAGCCAACGGTGAGGTGCT
ATGCACAGTGTGAGCGTGTCCCCAGACGGAGTGTGTGGACCCTGTGTACGAGCCTGATCAGTG
CTGTCCCATCTGCAAAAATGGTCCAAACTGCTTTGCAGAAACCGCGGTGATCCCTGCTGGCAG
AGAAGTGAAGACTGACGAGTGCACCAATATGCCACTGTACTTATGAGGAAGGCACATGGAGAAT
CGAGCGGCAGGCCATGTGCACGAGACATGAATGCAGGCAAATGTAGACGCTTCCCAGAACACA
AACTCTGACTTTTTCTAGAACATTTTACTGATGTGAACATTCTAGATGACTCTGGGAACTATC
AGTCAAAGAAGACTTTTGATGAGGAATAATGGAAAATTGTTGGTACTTTTCCTTTTCTTGATA
ACAGTTACTACAACAGAAGGAAATGGATATATTTCAAACATCAACAAGAACTTTGGGCATAA
AATCCTTCTCTAAATAAATGTGCTATTTTCACAGTAAGTACACAAAAGTACACTATTATATAT
CAAATGTATTTCTATAATCCCTCCATTAGAGAGCTTATATAAGTGTTTCTATAGATGCAGAT
TAAAAATGCTGTGTTGTCAACCGTCAAAAAAAAAAAAAAAAAAAAAAAAAA

516/550

FIGURE 516

MPSSTAMAVGALSSSLLVTCCLMVALCSPSIPLEKLAQAPEQPGQEKREHATRDGPGRVNELG
RPARDEGGSGRDWKS KSGRGLAGREPWSKLKQAWVSQGGGAKAGDLQVRPRGDT PQAEALAAA
AQDAIGPELAPTPEPPEEYVYPDYRGKGCVDSESGFVYAIGKFAFGPSACPCLCTEEGPLCAQ
PECPRLHPRCIHVDTSQCCPQCKERKNYCEFRGKTYQTLEEFVVSPCERCRCCEANGEVLCTVS
ACPQTECVDPVYEPDQCCPICKNGPNCFAETA V I PAGREVKTDECTICHCTYEETWRIERQA
MCTRHECRQM

Important features:**Signal peptide:**

amino acids 1-27

Transmembrane domain:

amino acids 11-30

Glycosaminoglycan attachment site.

amino acids 80-83

N-myristoylation sites.

amino acids 10-15, 102-107, 103-108

Cell attachment sequence.

amino acids 114-117

EGF-like domain cysteine pattern signature.

amino acids 176-187

517/550

FIGURE 517

GGACAACCGTTGCTGGGTGTCCCAGGGCCTGAGGCAGGACGGTACTCCGCTGACACCTTCCTT
TTCGGCCTTGAGGTTCCCAGCCTGGTGGCCCCAGGACGTTCCGGTCGCATGGCAGAGTGCTAC
GGACGACGCCTATGAAAGCCCTTAGTCCTTCTAGTTGCGCTTTTGCTATGGCCTTCGTCTGTGC
CGGCTTATCCGAGCATAACTGTGACACCTGATGAAGAGCAAACTTGAATCATTATATACAAG
TTTTAGAGAACCTAGTACGAAGTGTTCCCTCTGGGGAGCCAGGTCGTGAGAAAAATCTAACT
CTCCAAAACATGTTTATTCTATAGCATCAAAGGGATCAAAATTTAAGGAGCTAGTTACACATG
GAGACGCTTCAACTGAGAATGATGTTTTAACCAATCCTATCAGTGAAGAACTACAACCTTCC
CTACAGGAGGCTTCACACCGGAAATAGGAAAGAAAAACACACGGAAAGTACCCCATTTCTGGT
CGATCAAACCAAACAATGTTTCCATTGTTTTGCATGCAGAGGAACCTTATATTGAAAATGAAG
AGCCAGAGCCAGAGCCGGAGCCAGCTGCAAAACAACTGAGGCACCAAGAATGTTGCCAGTTG
TTACTGAATCATCTACAAGTCCATATGTTACCTCATACAAGTCACCTGTCACCACTTTAGATA
AGAGCACTGGCATTGAGATCTCTACAGAATCAGAAGATGTTCCCTCAGCTCTCAGGTGAACTG
CGATAGAAAAACCCGAAGAGTTTGGAAGCACCCAGAGAGTTGGAATAATGATGACATTTTGA
AAAAAATTTTAGATATTAATTCACAAGTGCAACAGGCACTTCTTAGTGACACCAGCAACCCAG
CATATAGAGAAGATATTGAAGCCTCTAAAGATCACCTAAAACGAAGCCTTGCTCTAGCAGCAG
CAGCAGAACATAAATTA AAAACAATGTATAAGTCCCAGTTATTGCCAGTAGGACGAACAAGTA
ATAAAATTGATGACATCGAACTGTTATTAACATGCTGTGTAATTCTAGATCTAACTCTATG
AATATTTAGATATTAAATGTGTTCCACCAGAGATGAGAGAAAAAGCTGCTACAGTATTCAATA
CATTA AAAAATATGTGTAGATCAAGGAGAGTCACAGCCTTATTA AAAGTTTATTAAACAATAA
TATAAAAATTTTAAACCTACTTGATATTCATAACAAAGCTGATTTAAGCAAAGCTGCATTTTT
TCACAGGAGAAATAATCATATTCGTAATTTCAAAGTTGTATAAAAATATTTTCTATTGTAGT
TCAAATGTGCCAACATCTTTATGTGTCATGTGTTATGAACAATTTTCATATGCACTAAAAACC
TAATTTAAATAAAAATTTTGGTTCAGGAAAAA

518/550

FIGURE 518

MKPLVLLVALLLWPSSVPAYPSITVTPDEEQNLNHYIQVLENLVRSVPSGEPGREKKSNSPKH
VYSIASKGSKFKELVTHGDASTENDVLNPISEETTFPTGGFTPEIGKKKHTESTPFWSIKP
NNVSIVLHAEOPYIENEPEPEPEPAKQTEAPRMLPVVTSSTSPYVTSYKSPVTTLDKSTG
IEISTESEDVPQLSGETAIEKPEEFGKHPESWNNDDILKKILDINSQVQQALLSDTSNPAYRE
DIEASKDHLKRSLALAAAAEHKLKTMYSQLLPVGRTSNKIDDIETVINMLCNSRSKLYEYLD
IKCVPPEMREKAATVFNTLKNMCRSRRVTALLKVY

Important features:**Signal peptide:**

amino acids 1-19

519/550

FIGURE 519

CGGCTCGAGTGCAGCTGTGGGGAGATTTTCAGTGCATTGCCTCCCCTGGGTGCTCTTCATCTTG
GATTTGAAAGTTGAGAGCAGC**ATG**TTTTGCCCACTGAAACTCATCCTGCTGCCAGTGTTACTG
GATTATTCCTTGGGCCTGAATGACTTGAATGTTTCCCCGCCTGAGCTAACAGTCCATGTGGGT
GATTCAGCTCTGATGGGATGTGTTTTCCAGAGCACAGAAGACAAATGTATATTTCAAGATAGAC
TGGACTCTGTCACCAGGAGAGCACGCCAAGGACGAATATGTGCTATACTATTACTCCAATCTC
AGTGTGCCTATTGGGCGCTTCCAGAACCGCGTACACTTGATGGGGGACATCTTATGCAATGAT
GGCTCTCTCCTGCTCCAAGATGTGCAAGAGGCTGACCAGGGAACCTATATCTGTGAAATCCGC
CTCAAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGGTACTGCATGTGCTTCCAGAGGAGCCC
AAAGAGCTCATGGTCCATGTGGGTGGATTGATTCAGATGGGATGTGTTTTCCAGAGCACAGAA
GTGAAACACGTGACCAAGGTAGAATGGATATTTTCAGGACGGCGCGCAAAGGAGGAGATTGTA
TTTCGTTACTACCACAACTCAGGATGTCTGTGGAGTACTCCCAGAGCTGGGGCCACTTCCAG
AATCGTGTGAACCTGGTGGGGGACATTTTCCGCAATGACGGTTCATCATGCTTCAAGGAGTG
AGGGAGTCAGATGGAGGAACTACACCTGCAGTATCCACCTAGGGAACTGGTGTTCAGAAA
ACCATTGTGCTGCATGTCAGCCCGGAAGAGCCTCGAACACTGGTGACCCCGGCAGCCCTGAGG
CCTCTGGTCTTGGGTGGTAATCAGTTGGTGATCATTGTGGGAATTGTCTGTGCCACAATCCTG
CTGCTCCCTGTTCTGATATTGATCGTGAAGAAGACCTGTGGAAATAAGAGTTCAGTGAATTCT
ACAGTCTTGGTGAAGAACACGAAGAAGACTAATCCAGAGATAAAAGAAAAACCCTGCCATTTT
GAAAGATGTGAAGGGGAGAAACACATTTACTCCCCAATAATTGTACGGGAGGTGATCGAGGAA
GAAGAACCAAGTGA AAAATCAGAGGCCACCTACATGACCATGCACCCAGTTTGGCCTTCTCTG
AGGTGAGATCGGAACAACCTCACTTGAAAAAAGTCAGGTGGGGGAATGCCAAAAACACAGCAA
GCCTTT**TGA**GAAGAATGGAGAGTCCCTTCATCTCAGCAGCGGTGGAGACTCTCTCCTGTGTGT
GTCCTGGGCCACTCTACCAGTGATTTTCAGACTCCCGCTCTCCAGCTGTCCTCCTGTCTCATT
GTTTGGTCAATACACTGAAGATGGAGAATTTGGAGCCTGGCAGAGAGACTGGACAGCTCTGGA
GGAACAGGCCTGCTGAGGGGAGGGGAGCATGGACTTGGCCTCTGGAGTGGGACACTGGCCCTG
GGAACCAGGCTGAGCTGAGTGGCCTCAAACCCCCGTTGGATCAGACCCTCCTGTGGGCAGGG
TTCTTAGTGGATGAGTTACTGGGAAGAATCAGAGATAAAACCAACCCAAATCAA

520/550

FIGURE 520

MFCPLKLILLPVLLDYSLGLNDLNVSPPELTVHVGDSALMGCVFQSTEDKCI FKIDWTLS PGE
HAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGTYICEIRLKGESQV
FKKAVVLHVLPEEPKELMVHVGGLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEEIVFRYYHKL
RMSVEYSQSWGHEFQNRVNLVGDI FRNDGSIMLQGVRES DGGNYTCSIHLGNLVFKKTIVLHVS
PEEPRTLVT PAALRPLVLGGNQLVIIVGIVCATILLPV LILIVKKT CGNKSSVNSTVLVKNT
KKTNP EIKEK PCHFERCEGEKHIYSPIIVREVIEEEEPSEKSEATYMTMHPVWPSLRSDRNNS
LEKKSGGGMPKTQQAF

521/550

FIGURE 521

CTATGAAGAAGCTTCCTGGAAAACAATAAGCAAAGGAAAACAAATGTGTCCCATCTCACATGG
TTCTACCCCTACTAAAGACAGGAAGATCATAAACTGACAGATACTGAAATTGTAAGAGTTGGAA
ACTACATTTTGCAAAGTCATTGAACTCTGAGCTCAGTTGCAGTACTCGGGAAGCCATGCAGGA
TGAAGATGGATACATCACCTTAAATATTAAAACTCGGAAACCAGCTCTCGTCTCCGTTGGCCC
TGCATCCTCCTCCTGGTGGCGTGTGATGGCTTTGATTCTGCTGATCCTGTGCGTGGGGATGGT
TGTCGGGCTGGTGGCTCTGGGGATTTGGTCTGTCATGCAGCGCAATTACCTACAAGATGAGAA
TGAAAATCGCACAGGAAGCTCTGCAACAATTAGCAAAGCGCTTCTGTCAATATGTGGTAAAACA
ATCAGAACTAAAGGGCACTTTCAAAGGTCATAAATGCAGCCCCTGTGACACAACTGGAGATA
TTATGGAGATAGCTGCTATGGGTTCTTCAGGCACAACTTAACATGGGAAGAGAGTAAGCAGTA
CTGCACTGACATGAATGCTACTCTCCTGAAGATTGACAACCGGAACATTGTGGAGTACATCAA
AGCCAGGACTCATTTAATTCGTTGGGTCGGATTATCTCGCCAGAAGTCGAATGAGGTCTGGAA
GTGGGAGGATGGCTCGGTTATCTCAGAAAATATGTTTGAGTTTTTGGAAGATGGAAAAGGAAA
TATGAATTGTGCTTATTTTCATAATGGGAAAATGCACCCTACCTTCTGTGAGAACAAACATTA
TTAATGTGTGAGAGGAAGGCTGGCATGACCAAGGTGGACCAACTACCTTAATGCAAAGAGGT
GGACAGGATAACACAGATAAGGGCTTTATTGTACAATAAAAGATATGTATGAATGCATCAGTA
GCTGAAAAAAAAAAAAA

522/550

FIGURE 522

MQDEGGYITLNIKTRKPALVSVGPASSSSWVRVMALILLILCVGMVVGLVALGIWSVMQARNYLQ
DENENRTGTLQQLAKRFCQYVVVKQSELKGTFKGHKCSPCDTNWRYYGDSYGFRRHNLTWEEES
KQYCTDMNATLLKIDNRNIVEYIKARTHLIRWVGLSRQKSNEVWKWEDGSVISENMFEFLEDG
KGNMNCAYFHNGKMHPTFCENKHLYLMCERKAGMTKVDQLP

523/550

FIGURE 523

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGC
AAAGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAA
GAAAATATGTAAATCACTTAAGATTTGTGGACTGGTGTGTTGGTATCCTGGCCCTAACTCTAAT
TGTCTGTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAAGCCTATGACATGGA
GCACACTTTCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTGACCAG
AACTGAAATATTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAA
CGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGT
GATTCCTGAATTTTCTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTT
CTTTGAACAGTCAGTGATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTTAA
AAATTCCAAAATTCTGGAGATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAAT
ATCAGTTTCTGAGTTACAAGACTTTGAGGAGGAGGGAGAAGATCTTCACTTTCCTGCCAACGA
AAAAAAGGGATTGAACAAAATGAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCG
TCACGCCAGACAAGCAAGTGAGGAAGAACTTCCAATAAATGACTATACTGAAAATGGAATAGA
ATTTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTATTTACTGCCGTCGAGGCAACCGCTA
TTGCCGCCGCGTCTGTGAACCTTTACTAGGCTACTACCCATATCCATACTGCTACCAAGGAGG
ACGAGTCATCTGTCGTGTCATCATGCCTTGTAAGTGGTGGGTGGCCCGCATGCTGGGGAGGGT
CTAATAGGAGGTTTGAGCTCAAATGCTTAACTGCTGGCAACATATAATAAATGCATGCTATT
CAATGAATTTCTGCCTATGAGGCATCTGGCCCCTGGTAGCCAGCTCTCCAGAATTACTTGTAG
GTAATTCCTCTCTTCATGTTCTAATAAACTTCTACATTATCACCAAAAAAAAAAAAAAAAAA

524/550

FIGURE 524

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALT LIVLFWGSKHFWPEVPPKKAYD
MEHTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKTQI
KVIPEFSEPEEEIDENEEITTTFFEQSVIWPVPAEKPIENRDFLKN SKILEICDNVTMYWINPT
LISVSELQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENG
IEFDPMLDERGYCCIIYCRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features:**Signal peptide:**

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

525/550

FIGURE 525

AGTGACAATCTCAGAGCAGCTTCTACACCACAGCCATTTCCAGCATGAAAGATCACTGGGGGTC
TCCTTCTGCTCTGTACAGTGGTCTATTTCTGTAGCAGCTCAGAAGCTGCTAGTCTGTCTCCAA
AAAAAGTGGACTGCAGCATTTACAAGAAGTATCCAGTGGTGGCCATCCCCTGCCCCATCACAT
ACCTACCAGTTTGTGGTTCTGACTACATCACCTATGGGAATGAATGTCACCTTGTGTACCGAGA
GCTTGAAAAGTAATGGAAGAGTTCAGTTTCTTCACGATGGAAGTTGCTAAATTCTCCATGGAC
ATAGAGAGAAAGGAATGATATTCTCATCATCATCTTCATCATCCCAGGCTCTGACTGAGTTTC
TTTCAGTTTTACTGATGTTCTGGGTGGGGGACAGAGCCAGATTCAGAGTAATCTTGAATGAAT
GGAGAAAGTTTCTGTGCTACCCCTACAAACCCATGCCTCACTGACAGACCAGCATTTTTTTTTT
TAACACGTCAATAAAAAAATAATCTCCCAGA

526/550

FIGURE 526

MKITGGLLLLCTVVYFCSSSEAASLSPKKVDCSIYKKYPVVAIPCPITYLPVCGSDYITYGNE
CHLCTESLKSNGRVQFLHDGSC

Important features:

Signal peptide:

amino acids 1-19

527/550

FIGURE 527

CGACG**ATG**CTACGCGCGCCCGGCTGCCTCCTCCGGACCTCCGTAGCGCCTGCCGCGGGCCCTGG
CTGCGGCGCTGCTCTCGTCGCTTGCGCGCTGCTCTCTTAGAGCCGAGGGACCCGGTGGCCT
CGTCGCTCAGCCCCTATTTTCGGCACCAAGACTCGCTACGAGGATGTCAACCCCGTGCTATTGT
CGGGCCCCGAGGCTCCGTGGCGGGACCCTGAGCTGCTGGAGGGGACCTGCACCCCGGTGCAGC
TGGTCGCCCCTCATTGCCACGGCACCCGCTACCCACGGTCAAACAGATCCGCAAGCTGAGGC
AGCTGCACGGGTTGCTGCAGGCCCGCGGGTCCAGGGATGGCGGGGCTAGTAGTACCGGCAGCC
GCGACCTGGGTGCAGCGCTGGCCGACTGGCCTTTGTGGTACGCGGACTGGATGGACGGGCAGC
TAGTAGAGAAGGGACGGCAGGATATGCGACAGCTGGCGCTGCGTCTGGCCTCGCTCTTCCCGG
CCCTTTTCAGCCGTGAGAACTACGGCCGCTGCGGCTCATCACCAGTTCCAAGCACCGCTGCA
TGGATAGCAGCGCCGCTTCCCTGCAGGGCTGTGGCAGCACTACCACCCTGGCTTGCCGCGC
CGGACGTGCGCAGATATGGAGTTTGGACCTCCAACAGTTAATGATAAACTAATGAGATTTTTTG
ATCACTGTGAGAAGTTTTTAAGTGAAGTAGAAAAAATGCTACAGCTCTTTATCACGTGGAAG
CCTTCAAACTGGACCAGAAATGCAGAACATTTTTAAAAAAGTTGCAGCTACTTTGCAAGTGC
CAGTAAATGATTTAAATGCAGATTTAATTCAAGTAGCCTTTTTTACCTGTTTCAATTTGACCTGG
CAATTAAGGTGTTAAATCTCCTTGGTGTGATGTTTTTGACATAGATGATGCAAAGGTATTAG
AATATTTAAATGATCTGAAACAATATTGAAAAGAGGATATGGGTATACTATTAACAGTCGAT
CCAGCTGCACCTTGTTCAGGATATCTTTCAGCACTTGGACAAAGCAGTTGAACAGAAACAAA
GGTCTCAGCCAATTTCTTCTCCAGTCATCTCCAGTTTGGTCATGCAGAGACTCTTCTTCCAC
TGCTTTCTCTCATGGGCTACTTCAAAGACAAGGAACCCCTAACAGCGTACAATTACAAAAAAC
AAATGCATCGGAAGTTCCGAAGTGGTCTCATTGTACCTTATGCCTCGAACCTGATATTTGTGC
TTTACCCTGTGAAAATGCTAAGACTCCTAAAGAACAATTCCGAGTGCAGATGTTATTAAATG
AAAAGGTGTTACCTTTGGCTTACTCACAAGAACTGTTTTCATTTTATGAAGATCTGAAGAACC
ACTACAAGGACATCCTTCAGAGTTGTCAAACCAGTGAAGAATGTGAATTAGCAAGGGCTAACA
GTACATCTGATGAAGTAT**TGAG**TAACTGAAGAACATTTTTAATTCTTTAGGAATCTGCAATGAG
TGATTACATGCTTGTAATAGGTAGGCAATTCCTTGATTACAGGAAGCTTTTATATTACTTGAG
TATTTCTGTCTTTTCACAGAAAAACATTGGGTTTCTCTCTGGGTTTGGACATGAAATGTAAGA
AAAGATTTTTTCACTGGAGCAGCTCTCTTAAGGAGAAACAAATCTATTTAGAGAAACAGCTGGC
CCTGCAAATGTTTACAGAAATGAAATTCCTTCTACTTATATAAGAAATCTCACACTGAGATAG
AATTGTGATTTTATAATAACACTTGAAAAGTGCTGGAGTAACAAAATATCTCAGTTGGACCAT
CCTTAACCTGATTGAAGTGTCTAGGAACCTTTACAGATTGTTCTGCAGTTCTCTCTTCTTTCC
TCAGGTAGGACAGCTCTAGCATTTTCTTAATCAGGAATATTGTGGTAAGCTGGGAGTATCACT
CTGGAAGAAAGTAACATCTCCAGATGAGAATTTGAAACAAGAAACAGAGTGTTGTAAAAGGAC
ACCTTCACTGAAGCAAGTCGGAAAGTACAATGAAAATAAATATTTTGGTATTTATTTATGAA
ATATTTGAACATTTTTTCAATAATTCCTTTTACTTCTAGGAAGTCTCAAAGACCATCTTAA
ATTATTATATGTTTGGACAATTAGCAACAAGTCAGATAGTTAGAATCGAAGTTTTTCAAATCC
ATTGCTTAGCTAACTTTTTCATTTCTGTCACTTGGCTTCGATTTTTATTTTTCTATTATATG
AAATGTATCTTTTGGTTGTTTGATTTTCTTTCTTTCTTTGTAAATAGTTCTGAGTTCTGTCA
AATGCCGTGAAAGTATTTGCTATAATAAGAAAATTCTTGTGACTTTAAAAAAA

528/550

FIGURE 528

MLRAPGCLLRTSVAPAAALAAALLSSLARCSLLEPRDPVASSLSPLYFGTKTRYEDVNPVLLSG
PEAPWRDPELLEGTCTPVQLVALIRHGTRYPTVKQIRKLRQLHGLLQARGSRDGGASSTGSRD
LGAALADWPLWYADWMDGQLVEKGRQDMRQLALRLASLFPALFSRENYGRLRLITSSKHRCMD
SSAFLQGLWQHYHPGLPPPDVADMEFGPPTVNDKLMRFFDHCEKFLTEVEKNATALYHVEAF
KTGPEMQNILKKVAATLQVPVNDLNADLIQVAFFTCSEFDLAIKGVKSPWCDVFDIDDAKVLEY
LNDLKQYWKRGYGYTINSRSSCTLFQDIFQHLDKAVEQKQRSQPISSPVILQFGHAETLLPLL
SLMGYFKDKEPLTAYNYKKQMHRKFRSGLIVPYASNLI FVLYHCENAKTPKEQFRVQMLLNEK
VLPLAYSQETVSFYEDLKNHYKDILQSCQTSEECELARANSTSDEL

Important features:**Signal sequence**

amino acids 1-30

N-glycosylation sites.

amino acids 242-246, 481-485

N-myristoylation sites.

amino acids 107-113, 113-119, 117-123, 118-124, 128-134

Endoplasmic reticulum targeting sequence.

amino acids 484-489

529/550

FIGURE 529

GGAGAGCCGCGGCTGGGACCGGAGTGGGGAGCGCGGCGTGGAGGTGCCACCCGGCGCGGGTGG
CGGAGAGATCAGAAGCCTCTTCCCCAAGCCGAGCCAACCTCAGCGGGGACCCGGGCTCAGGGA
CGCGGCGGCGGCGGCGGCGGCGACTGCAGTGGCTGGACGATGGCAGCGTCCGCCGGAGCCGGGGCG
GTGATTGCAGCCCCAGACAGCCGGCGCTGGCTGTGGTTCGGTGTGGCGGCGGCGCTTGGGCTC
TTGACAGCTGGAGTATCAGCCTTGGAAAGTATATACGCCAAAAGAAATCTTCGTGGCAAATGGT
ACACAAGGGAAGCTGACCTGCAAGTTCAAGTCTACTAGTACGACTGGCGGGTTGACCTCAGTC
TCCTGGAGCTTCCAGCCAGAGGGGGCCGACACTACTGTGTCGTTTTTCCACTACTCCCAAGGG
CAAGTGTACCTTGGGAATTATCCACCATTTAAAGACAGAATCAGCTGGGCTGGAGACCTTGAC
AAGAAAGATGCATCAATCAACATAGAAAATATGCAGTTTATACACAATGGCACCTATATCTGT
GATGTCAAAAACCTCCTGACATCGTTGTCCAGCCTGGACACATTAGGCTCTATGTCTGTAGAA
AAAGAGAATTTGCCTGTGTTTCCAGTTTGGGTAGTGGTGGGCATAGTTACTGCTGTGGTCTTA
GGTCTCACTCTGCTCATCAGCATGATTCTGGCTGTCTCTATAGAAGGAAAACTCTAAACGG
GATTACACTGGCTGCAGTACATCAGAGAGTTTGTCCACAGTTAAGCAGGCTCCTCGGAAGTCC
CCCTCCGACACTGAGGGTCTTGTAAAGAGTCTGCCTTCTGGATCTCACCAGGGCCAGTCATA
TATGCACAGTTAGACCACTCCGGCGGACATCACAGTGACAAGATTAACAAGTCAGAGTCTGTG
GTGTATGCGGATATCCGAAAGAATTAAGAGAATACCTAGAACATATCCTCAGCAAGAAACAAA
ACCAAACTGGACTCTCGTGCAGAAAATGTAGCCATTACCACATGTAGCCTTGGAGACCCAGG
CAAGGACAAGTACACGTGTACTCACAGAGGGAGAGAAAGATGTGTACAAAGGATATGTATAAA
TATTCTATTTAGTCATCCTGATATGAGGAGCCAGTGTTCATGATGAAAAGATGGTATGATTC
TACATATGTACCCATTGTCTTGCTGTTTTTGTACTTTCTTTTCAGGTCAATTTACAATTGGGAG
ATTTCAGAAACATTCCTTTCACCATCATTTAGAAATGGTTTGCCTTAATGGAGACAATAGCAG
ATCCTGTAGTATTTCCAGTAGACATGGCCTTTTAACTAAGGGCTTAAGACTGATTAGTCTTA
GCATTTACTGTAGTTGGAGGATGGAGATGCTATGATGGAAGCATACCCAGGGTGGCCTTTAGC
ACAGTATCAGTACCATTTATTTGTCTGCCGCTTTTAAAAAATACCCATTGGCTATGCCACTTG
AAAACAATTTGAGAAGTTTTTTTTGAAGTTTTTCTCACTAAAATATGGGGCAATTGTTAGCCTT
ACATGTTGTGTAGACTTACTTTAAGTTTTGCACCCTTGAAATGTGTCATATCAATTTCTGGATT
CATAATAGCAAGATTAGCAAAGGATAAATGCCGAAGGTCACTTCATTCTGGACACAGTTGGAT
CAATACTGATTAAGTAGAAAATCCAAGCTTTGCTTGAGAACTTTTGTAACGTGGAGAGTAAAA
AGTATCGGTTTTTA

530/550

FIGURE 530

MAASAGAGAVIAAPDSRRWLWSVLAAALGLLTAGVSALEVYTPKEIFVANGTQGKLTCKFKST
STTGGLTSVSWSFQPEGADTTVSFFHYSQGQVYLGNYPPFKDRISWAGDLDDKKDASINIENMQ
FIHNGTYICDVKNPPDIVVQPGHIRLYVVEKENLPVFPVWVVVGIVTAVVLGLTLLISMILAV
LYRRKNSKRDTGCTSESLSPVKQAPRKSPSDTEGLVKSLPSGSHQGPVIYAQLDHSGGHHS
DKINKSESVVYADIRKN

Important features:

Signal peptide:

amino acids 1-37

Transmembrane domain:

amino acids 161-183

531/550

FIGURE 531

GTGACACTATAGAAGAGCTATGACGTCGCATGCACGCGTACGTAAGCTCGGAATTCGGCTCGA
GGCTGGTGGGAAGAAGCCGAG**ATG**GCGGCAGCCAGCGCTGGGGCAACCCGGCTGCTCCTGCTC
TTGCTGATGGCGGTAGCAGCGCCCAGTCGAGCCCGGGGCAGCGGCTGCCGGGCCGGGACTGGT
GCGCGAGGGGGCTGGGGCGGAAGGTCGAGAGGGCGAGGCCCTGTGGCACGGTGGGGCTGCTGCTG
GAGCACTCATTTGAGATCGATGACAGTGCCAACTTCCGGAAGCGGGGCTCACTGCTCTGGAAC
CAGCAGGATGGTACCTTGTCCTGTACAGCGGCAGCTCAGCGAGGAGGAGCGGGGCCGACTC
CGGGATGTGGCAGCCCTGAATGGCCTGTACCGGGTCCGGATCCCAAGGCGACCCGGGGCCCTG
GATGGCCTGGAAGCTGGTGGCTATGTCTCCTCCTTTGTCCCTGCGTGCTCCCTGGTGGAGTCG
CACCTGTCGGACCAGCTGACCCTGCACGTGGATGTGGCCGGCAACGTGGTGGGCGTGTCGGTG
GTGACGCACCCCGGGGGCTGCCGGGGCCATGAGGTGGAGGACGTGGACCTGGAGCTGTTCAAC
ACCTCGGTGCAGCTGCAGCCGCCACCACAGCCCCAGGCCCTGAGACGGCGGCCTTCATTGAG
CGCCTGGAGATGGAACAGGCCCAGAAGGCCAAGAACCCCCAGGAGCAGAAGTCCTTCTTCGCC
AAATACTGGATGTACATCATTTCCCGTCGTCCTGTTCCCTCATGATGTCAGGAGCGCCAGACACC
GGGGGCCAGGGTGGGGGTGGGGGTGGGGGTGGTGGTGGGGGTAGTGGCCTTTGCTGTGTGCCA
CCCTCCCTG**TAAG**TCTATTTAAAAACATCGACGATACATTGAAATGTGTGAACGTTTTGAAAA
GCTACAGCTTCCAGCAGCCAAAAGCAACTGTTGTTTTGGCAAGACGGTCCTGATGTACAAGCT
TGATTGAAATTCAGTCTCACTTGATACGTTATTTCAGAAACCCAAGGAATGGCTGTCCCCATC
CTCATGTGGCTGTGTGGAGCTCAGCTGTGTTGTGTGGCAGTTTATTAACTGTCCCCAGATC
GACACGCAAAAAAAAAA

532/550

FIGURE 532

MAAASAGATRLLLLLLMAVAAPSRARGSGCRAGTGARGAGAEGREGEACGTVGLLLEHSFEID
DSANFRKRGSLLWNQQDGTLSLSQRQLSEEERGRLRDVAALNGLYRVRI PRRP GALDGLEAGG
YVSSFVPACSLVESHLSDQLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDLELFNTSVQLQP
PTTAPGPETA AFIERLEMEQAQKAKNPQEQKSFFAKYWMYIIIPVVLFLMMSGAPDTGGQGGGG
GGGGGGGSGLC CVPPSL

Important features:

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 226-243

534/550

FIGURE 534

MELALLCGLVVMAGVPIPIQGGILNLNKMVKQVTGKMPILSYWPYGCHCGLGGRGQPKDATDWC
CQTHDCCYDHLKTQCGGIYKDNNKSSIHCMDL SQRYCLMAVFNVIIYLEDSE

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 1-24

N-glycosylation site.

amino acids 86-89

N-myristoylation sites.

amino acids 20-25, 45-50

Phospholipase A2 histidine active site.

amino acids 63-70

535/550

FIGURE 535

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTG
AACTGGGTGCTCATCACGGGAAGTCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCC
CCAAATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTT
TTTTTAACCGCCCCCTCCCCACCCCCCAAAAAAAGTGTAAAGATGCAAAAACGTAATATCCAT
GAAGATCCTATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTT
GTTCTTGGAGTGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAG
GGGTCCAATTTTTCTTCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGG
CTGTCATGCAACTGGCCCCTAAGCCAAAGCAAAGACCTAAGGACGACCTTTGAACAATACAA
AGGATGGGTTTTCAATGTAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCC
ACTGTCTTACTGACAATGCTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAA
GGCAAAATGGTATATTGTGAATCTCAGAAATTACAGGAGATACCCTCAAGTATATCTGCTGGT
TGCTTAGGTTTGTCCCTTCGCTATAACAGCCTTCAAAAAGTAAAGTATAATCAATTTAAAGGG
CTCAACGAGTACCTACCTGGCTATACCTTGACCATAACCATATCAGCAATATTGACGAAAATGCT
TTTAATGGAATACGAGACTCAAAGAGCTGATTCTTAGTTCCAATAGAATCTCCTATTTTCTT
AACAATACCTTCAGACCTGTGACAAATTTACGGAAGCTTGATCTGTCTTATAATCAGCTGCAT
TCTCTGGGATCTGAACAGTTTCGGGGCTTGCGGAAGCTGCTGAGTTTACATTTACGGTCTAAC
TCCCTGAGAACCATCCCTGTGCGAATATTCCAAGACTGCCGCAACCTGGAACTTTTGGACCTG
GGATATAACCGGATCCGAAGTTTAGCCAGGAATGTCTTTGCTGGCATGATCAGACTCAAAGAA
CTTACCTGGAGCACAATCAATTTTCCAAGCTCAACCTGGCCCTTTTTCCAAGGTTGGTCAGC
CTTCAGAACCTTTACTTGCAGTGGAATAAAATCAGTGTATAGGACAGACCATGTCCTGGACC
TGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGATCGAAGCTTTCAGTGGACCCAGT
GTTTTCCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGATTCCAACAAGCTCACATTTATT
GGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAGTCTTGCTGGGAATATATGG
GAATGCAGCAGAAATATTTGCTCCCTTGTAAGTGGCTGAAAAGTTTTAAAGGTCTAAGGGAG
AATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTGATCGATGCAGTGAAG
AACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTTGATCTGGCCAGGGCTCTCCCAAAG
CCGACGTTTAAGCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTTTGCCCCCGACG
GTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCTTTCCATAAA
ATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCGTGCTCGTCATCCTGCTGGTTATCTACGTG
TCATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAAGGCAC
AGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCCAGCACCCAGGAATTTTATGTAGAT
TATAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCCTGCACCTAT
AACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTTGTGATAAAAAGAGCTCTTAAAGC
TGGGAAATAAGTGGTGCTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCCTC
CCCTTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCCTTGTCGTTTTAGTGCATTCTATA
ATACTGGTCATTTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGT
GAAGCTTGAATCCGGTTTAAATATAATACCTATTGTATAAGACCTTTACTGATTCCATTAAT
GTCGCATTTGTTTTAAGATAAAACTTCTTTCATAGGTAAAAA

536/550

FIGURE 536

MGFNVIRLLSGSAVALVIAPT VLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGC
LGLSLRYNSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLN
NTFRPVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRI FQDCRNLELLDLG
YNRIRSLARNVFAGMIRLKEHLEHNQFSKLNALFPRLVSLQONLYLQWNKISVIGQTMSWTW
SSLQRLDLSGNEIEAFSGPSVFQCVPNLQRLNLD SNKLTFIGQEILDSWISLNDISLAGNIWE
CSRNICSLVNWLKSFKGLRENTIICASPKELOGVNVIDAVKNYSICGKSTTERFDLARALPKP
TFKPKLPRPKHESKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVS
WKRYPASMKQLQQRSLMRRHRKKRQSLKQMPSTQEFYVDYKPTNTETSEMLLNGTGPCTYN
KSGSRECEV

Important features:**Signal peptide:**

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

537/550

FIGURE 537

GGGACTACAAGCCGCGCCGCGCTGCCGCTGGCCCCCTCAGCAACCCTCGACATGGCGCTGAGGCGGCCACCGCGAC
TCCGGCTCTGCGCTCGGCTGCCTGACTTCTTCTGCTGCTGCTTTTCAGGGGCTGCCTGATAGGGGCTGTAAATC
TCAAATCCAGCAATCGAACCCAGTGGTACAGGAATTTGAAAGTGGAACCTGCTTGCATCATTACGGATTCCG
AGACAAGTGACCCAGGATCGAGTGGAAGAAAATTCAAGATGAACAAACCACATATGTGTTTTTTGACAACAAAA
TTCAGGGAGACTTGGCGGCTCGTGCAGAAATACTGGGGAAGACATCCCTGAAGATCTGGAATGTGACACGGAGAG
ACTCAGCCCTTTATCGCTGTGAGGTCGTTGCTCGAAATGACCGCAAGGAAATTGATGAGATTGTGATCGAGTTAA
CTGTGCAAGTGAAGCCAGTGACCCCTGTCTGTAGAGTGCCGAAGGCTGTACCAGTAGGCAAGATGGCAACACTGC
ACTGCCAGGAGAGTGAGGGCCACCCCGGCCCTCACTACAGCTGGTATCGCAATGATGTACCACTGCCACGGATT
CCAGAGCCAATCCCAGATTTTCGCAATTCTTCTTTCCACTTAAACTCTGAAACAGGCACCTTGGTGTTCAGTGCTG
TTCACAAGGACGACTCTGGGCGTACTACTGCATTGCTTCCAATGACGCAGGCTCAGCCAGGTGTGAGGAGCAGG
AGATGGAAGTCTATGACCTGAACATTGGCGGAATTATTGGGGGGGTTCTGGTGTCTTGTCTACTGGCCCTGA
TCACGTTGGGCATCTGCTGTGCATACAGACGTGGCTACTTCATCAACAATAAACAGGATGGAGAAAGTTACAAGA
ACCCAGGGAACCCAGATGGAGTTAACTACATCCGACACTGACGAGGAGGCGACTTCAGACACAAGTCATCGTTTG
TGATCTGAGACCCGCGGTGTGGCTGAGAGCGCACAGAGCCGACGTCACATACCTCTGCTAGAAACTCCTGTCAA
GGCAGCGAGAGCTGATGCACTCGGACAGAGCTAGACACTCATTGAGAAGCTTTTCGTTTTGGCCAAAGTTGACCA
CTACTCTTCTTACTTAACAAGCCACATGAATAGAAGAAATTTTCTCAAGATGGACCCGGTAAATATAACCACAA
GGAAGCGAAACTGGGTGCGTTCACTGAGTTGGGTTCCTAATCTGTTTCTGGCCTGATTCCCGCATGAGTATTAGG
GTGATCTTAAAGAGTTTGTCTACGTAAACGCCGCTGCTGGGCCCTGTGAAGCCAGCATGTTCACTACTGGTCGTT
CAGCAGCCACGACAGCACCATGTGAGATGGCGAGGTGGCTGGACAGCACCAGCAGCGCATCCCGCGGGAACCCA
GAAAAGGCTTCTTACACAGCAGCCTTACTTCATCGGCCACAGACACCACCGCAGTTTCTTCTTAAAGGCTCTGC
TGATCGGTGTTGCAGTGTCCATTGTGGAGAAGCTTTTTGGATCAGCATTGTGTAACAAACCAAAATCAGGAAG
GTAAATTGGTTGCTGGAAGAGGGATCTTGCCCTGAGGAACCTGCTTGTCCAACAGGGTGTGAGGATTTAAGGAAA
ACCTTCGTTCTAGGCTAAGTCTGAAATGGTACTGAAATATGCTTTTCTATGGGTCTTGTTTTATTTTATAAAATTT
TACATCTAAATTTTTGCTAAGGATGTATTTTGTATTATGAAAAGAAAATTTCTATTAAACTGTAAATATATTGT
CATACAATGTTAAATAACCTATTTTTTTAAAAAAGTTCAACTTAAGGTAGAAGTTCCAAGCTACTAGTGTTAAAT
TGGAAAATATCAATAATTAAGAGTATTTTACCCAAGGAATCCTCTCATGGAAGTTTACTGTGATGTTCCTTTTCT
CACACAAGTTTTAGCCTTTTTTACAAGGGAACCTACTACTGTCTACACATCAGACCATAGTTGCTTAGGAAACCTT
TAAAAATTCAGTTAAGCAATGTTGAAATCAGTTTGCATCTCTTCAAAGAAACCTCTCAGGTAGCTTTGAACT
CCCTCTTCTGAGATGACTAGGACAGTCTGTACCCAGAGGCCACCCAGAAGCCCTCAGATGTACATACACAGATG
CCAGTCAGCTCCTGGGGTTGCGCCAGGCGCCCCGCTCTAGCTCACTGTTGCCTCGCTGTCTGCCAGGAGGCCCT
GCCATCCTTGGGCCCTGGCAGTGGCTGTGTCCCAGTGAGCTTTACTCACGTGGCCCTTGCTTCATCCAGCACAGC
TCTCAGGTGGGCACTGCAGGGCACTGGTGTCTTCCATGTAGCGTCCCAGCTTTGGGCTCCTGTAACAGACCTCT
TTTTGGTTATGGATGGCTCACAAAATAGGGCCCCCAATGCTATTTTTTTTTTTAAGTTTGTTAATTATTGTT
AAGATTGTCTAAGGCCAAAGGCAATTGCGAAATCAAGTCTGTCAAGTACAATAACATTTTTTAAAGAAAATGGAT
CCCCTGTTCTCTTGGCCACAGAGAAAGCACCCAGACGCCACAGGCTCTGTGCGATTTCAAACAAACCATGAT
GGAGTGGCGGCCAGTCCAGCCTTTTAAAGAACGTCAAGTGGAGCAGCCAGGTGAAAGGCCCTGGCGGGGAGGAAAG
TGAAACGCCCTGAATCAAAGCAGTTTTCTAATTTTGACTTTAAATTTTTCATCCGCCGGAGACACTGCTCCCATT
TGTGGGGGACATTAGCAACATCACTCAGAAGCCTGTGTTCTTCAAGAGCAGGTGTTCTCAGCCTCACATGCCCT
GCCGTGCTGGACTCAGGACTGAAGTGTGTAAAGCAAGGAGCTGCTGAGAAGGAGCACTCCACTGTGTGCTGGA
GAATGGCTCTCACTACTCACCTTGTCTTTAGCTTCCAGTGTCTTGGGTTTTTTTACTTTGACAGCTTTTTTTT
AATTGCATACATGAGACTGTGTTGACTTTTTTTTAGTTATGTGAAACACTTTGCCCGAGGCCCTGGCAGAGGCA
GGAAATGCTCCAGCAGTGGCTCAGTGTCTCCCTGGTGTCTGCTGCATGGCATCCTGGATGCTTAGCATGCAAGTTC
CCTCATCATTGCCACCTTGGTAGAGAGGGATGGTCCCCACCCTCAGCGTTGGGGATTACGCTCCAGCCTCCT
TCTTGGTTGTATAGTGATAGGGTAGCCTTATTGCCCCCTCTTCTTATACCCTAAAACCTTCTACACTAGTGCCA
TGGGAACCAGGTCTGAAAAAGTAGAGAGAAGTAAAGTAGAGTCTGGGAAGTAGCTGCCTATAACTGAGACTAGA
CGGAAAAGGAATACTCGTGTATTTTAAAGATATGAATGTGACTCAAGACTCGAGGCCGATACGAGGCTGTGATTCT
GCCTTTGGATGGATGTTGCTGTACACAGATGCTACAGACTGTACTAACACACCGTAATTTGGCATTGTGTTAAC
CTCATTATATAAAGCTTCAAAAAACCCA

538/550

FIGURE 538

MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTPVVQEFESVELSCIITDSQTSDF
RIEWKKIQDEQTTYVFFDNKIQGDLAGRAEILGKTSLSKIWNVTRRDSALYRCEVVARNDRKEI
DEIVIELTVQVKPVPVCRVPKAVPVGKMATLHCQESEGHPRPHYSWYRNDVPLPTDSRANPR
FRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVDNLNIGGIIGGVLVV
LAVLALITLGICCAYYRRGYFINNKQDGESYKNPGKPDGVNYIRTDEEGDFRHKSSFVI

Important features:

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 243-263

N-glycosylation sites.

amino acids 104-107, 192-195

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 107-110

Casein kinase II phosphorylation site.

amino acids 106-109, 296-299

Tyrosine kinase phosphorylation site.

amino acids 69-77

N-myristoylation sites.

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267

539/550

FIGURE 539

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCTCAGAGGCCGGGGAAGAGAAGCAAAGC
GCAACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCCCTAA
CTTCAGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCGCGG
CACAGGCGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTCTGGGCAGCTGG
GCTCGGGCGGGCGGGAGTAGGGCCCCGGCAGGGAGGCGAGGGAGGCTGCATATTCAGAGTCGCGGG
CTGCGCCCTGGGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGA
TGAGCCGCGTGGTCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCC
GCCGCGTGGTCAAGCGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGG
CCTACTTCCATGAAGTGTCCAGCCGAGTGAGCTTTTCAGGAGGCACGCCTGGCTTGTGAGAGTG
AGGGAGGAGTCCCTCCTCAGCCTTGAGAATGAAGCAGAACAGAAAGTTAATAGAGAGCATGTTGC
AAAACCTGACAAAACCCGGGACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGA
ATGGAGATGGGCAAACATCTGGTGCCTGCCCAGATCTCTACCAGTGGTCTGATGGAAGCAATT
CCCAGTACCGAACTGGTACACAGATGAACCTTCCTGCGGAAGTGAAAAGTGTGTTGTGATGT
ATCACCACCAACTGCCAATCCTGGCCTTGGGGGTCCCTACCTTTACCAGTGGAATGATGACA
GGTGTAAACATGAAGCACAAATTATATTTGCAAGTATGAACCAGAGATTAATCCAACAGCCCCCTG
TAGAAAAGCCTTATCTTACAAATCAACCAGGAGACACCCATCAGAATGTGGTTGTTACTGAAG
CAGGTATAATTCCCAATCTAATTTATGTTGTTATACCAACAATACCCCTGCTCTTACTGATAC
TGGTTGCTTTTGGAACCTGTTGTTTCCAGATGCTGCATAAAAAGTAAAGGAAGAACAACAACTA
GTCCAAACCAGTCTACACTGTGGATTTCAAAGAGTACCAGAAAAGAAAGTGGCATGGAAGTAT
AATAACTCATTGACTTGGTTCCAGAATTTTGTAATTTCTGGATCTGTATAAGGAATGGCATCAG
ACAATAGCCTTGGAATGGCTTGAAATCACAAAGGATCTGCAAGATGAAGCTGAAGCTCCCCCT
TGAGGCAATATTAAAGTAATTTTATATGTCTATTATTTTCAATTAAGAATATGCTGTGCTA
ATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAACTTCAAACCTCAAGCAAA
TGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTATGTGTGTTAGAAGCAAT
TCCTTTTATTTCTTTTACCTTTTATAAGTTGTTATCTAGTCAATGTAATGTATATTGTATTGA
AATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGATAAAAATGAACTGTTCTA
ATATTTATTTTATGGCATCTCATTTTTCAATACATGCTCTTTTGATTAAAGAACTTATTAC
TGTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAGTTTGTCTTCTCGAA
ATAATTCATCTTTCAGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGAAATAAGA
AGCTATTTTATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGATTGTCT
AATTTCAATTGTGCAAGACATGTGCCTTATAATTATTTTGTAGCTTAAATTAACAGATTTTG
TAATAATGTAAGTTTGTAAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAAGT
GACATACACAATATAAATCATATGTCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTCTG
AGGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTCTGGGGTTT
GGGATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGC
CTCTGACTATATTAGTATACAAAAGAGGTCAATGTGGTTGAGACCAGGTGAATAGTCACTATCAG
TGTGGAGACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAAGTACGAAATCGTGTGAAA
ATGGGTTGGAACCCATCAGTGATCGCATATTCATTGATGAGGGTTTGCTTGAGATAGAAAATG
GTGGCTCCTTTCTGTCTTATCTCTAGTTTCTTCAATGCTTACGCCTTGTTCTTCTCAAGAGA
AAGTTGTAAGTCTCTGGTCTTCATATGTCCCTGTGCTCCTTTTAAACCAATAAAGAGTTCTTG
TTTCTGGGGGAAA

540/550

FIGURE 540

MSRVVSLLLGAALLCGHGAFRRVVSQGKVCFADFKHPCYKMAYFHELSSRVSFQEARLACES
EGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDGFWIGLWRNGDGQTSACPDLYQWSDGSN
SQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNKYCKYEPEINPTAP
VEKPYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSCKGRTKT
SPNQSTLWISKSTRKESGMEV

Important features:**Signal peptide:**

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89 and 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145
and 212-217

541/550

FIGURE 541

GGAGAAATGGAGAGAGCAGTGAAGTGGAGTCCGGGGTCTGGTCCGGGTGGTCTGTCTGCTCCTGGCATGCCCTG
CCACAGCCACTGGGCCCCGAAGTTGCTCAGCCTGAAGTAGACACCACCTGGGTCGTGTGCGAGGCCGCGAGGTGG
GCGTGAAGGGCACAGACCGCCTTGTGAATGTCTTTCTGGGCATTCCATTTGCCAGCCGCCACTGGGCCCTGACC
GGTTCTCAGCCCCACACCAGCACAGCCCTGGGAGGGTGTGCGGGATGCCAGCACTGCGCCCCAATGTGCCTAC
AAGACGTGGAGAGCATGAACAGCAGCAGATTTGTCCTCAACGAAAACAGCAGATCTTCTCCGTTTCAGAGGACT
GCCTGGTCTCAACGTCTATAGCCAGCTGAGGTCCCCGAGGGTCCGGTAGGCCGGTCATGGTATGGGTCCATG
GAGGCGCTCTGATAACTGGCGCTGCCACCTCCTACGATGGATCAGCTCTGGCTGCCTATGGGGATGTGGTCTGTG
TTACAGTCCAGTACCGCCTTGGGGTCTTGGCTTCTTCAGCACTGGAGATGAGCATGCACCTGGCAACAGGGCT
TCCTAGATGTGGTAGCTGCTTTGCGCTGGGTGCAAGAAAACATCGCCCCCTTCGGGGGTGACCTCAACTGTGTCA
CTGTCTTTGGTGGATCTGCCGTGGGAGCATCATCTCTGGCCTGGTCTGTCCCAGTGGCTGCAGGGCTGTTCC
ACAGAGCCATCACACAGAGTGGGGTCAACACCCAGGGATCATCGACTCTCACCTTGGCCCCTAGCTCAGA
AAATCGCAAAACCTTGGCCTGCAGCTCCAGCTCCCCGGCTGAGATGGTGCAGTGCCTTCAGCAGAAAGAAGGAG
AAGAGCTGGTCTTAGCAAGAAGCTGAAAAATACTATCTATCCTCTCACCGTTGATGGCACTGTCTCCCCAAAA
GCCCCAAGGAACCTCTGAAGGAGAAGCCCTTCCACTCTGTGCCCTTCTCATGGGTGTCAACAACCATGAGTTCA
GCTGGCTCATCCCCAGGGGTGGGGTCTCTGGATAACAATGGAGCAGATGAGCCGGGAGGACATGTGCCCATCT
CAACACCCGTCTTGACAGTCTGGATGTCCCCCTGAGATGATGCCACCGTCATAGATGAATACCTAGGAAGCA
ACTCGGACGCACAAGCCAAATGCCAGGCGTTCCAGGAATTCATGGGTGACGTATTCATCAATGTTCCACCGTCA
GTTTTTCAAGATACCTTCGAGATTCTGGAAGCCCTGTCTTTTCTATGAGTTCAGCATCGACCCAGTTCTTTTG
CGAAGATCAAACCTGCCTGGGTGAAGGCTGATCATGGGGCCGAGGGTGCTTTTGTGTTGCGAGGTCCCTTCTCA
TGGACGAGAGCTCCCGCTGGCCTTTCCAGAGGCCACAGAGGAGGAGAAGCAGCTAAGCCTCACCATGATGGCCC
AGTGGACCCACTTTGCCCGGACAGGGGACCCCAATAGCAAGGCTCTGCCTCCTTGGCCCCAATTCAACCAGGCGG
ACAATATCTGGAGATCAACCCAGTGCCACGGGCCGGACAGAAGTTCAGGGAGGCTGGATGCAGTCTGGTTCAG
AGACGCTCCCCAGCAAGATACAACAGTGGCACCAGAGAAGCAGAAGAACAGGAAGGCCAGGAGGACCTCTGAGGCC
AGGCTGAACCTTCTTGGCTGGGGCAAACCACTCTTCAAGTGGTGGCAGAGTCCCAGCACGGCAGCCCGCTCTC
CCCCTGCTGAGACTTTAATCTCCACCAGCCCTTAAAGTGTGCGCCGCTCTGTGACTGGAGTTATGCTCTTTTGAA
ATGTCACAAGGCCGCTCCACCTCTGGGGCATTGTACAAGTTCTTCCCTCTCCCTGAAGTGCCTTTCCTGCTTT
CTTCGTGGTAGGTTCTAGCACATTCTCTAGCTTCTTGGAGGACTCACTCCCCAGGAAGCCTTCCCTGCCTTCTC
TGGGCTGTGCGGCCCCGAGTCTGCGTCCATTAGAGCAGTCCACCCGAGGCTAGCACCGTGTCTGTCTGTCT
CCCCCTCAGAGGAGCTCTCTCAAATGGGGATTAGCCTAACCCCACTCTGTCAACCCACACCAGGATCGGGTGGGA
CCTGGAGCTAGGGGTGTTTGTGAGTGAAGTGAACACAGAATATGGGAATGGCAGCTGTGAACCTGAAC
CCAGAGCCTTCAGTGCCAAAGCCATACTCAGGCCCCACCGACATTGTCCACCTGGCCAGAAAGGTGCATGCC
AATGGCAGAGACCTGGGATGGGAGAAGTCTGGGGCGCCAGGGGATCCAGCCTAGAGCAGACCTTAGCCCCCTGAC
TAAGGCCTCAGACTAGGGCGGGAGGGGTCTCCTCCTCTGTCTGCCAGTCTTGGCCCCCTGCACAAGACAACAGA
ATCCATCAGGGCCATGAGTGTACCCAGACCTGACCTCACCAATTCCAGCCCCCTGACCTCAGGACGCTGGATG
CCAGCTCCCAGCCCCAGTGCCGGTCTCCTCTCCTTCTTGGCTTGGGGAGACAGTTTCTGGGGAGCTTCCAAG
AGCACCCACCAAGACACAGCAGGACAGGCCAGGGGAGGGCATCTGGACCAGGGCATCCGTGGGGCTATTGTACA
GAGAAAAGAAGAGACCCACCACTCGGGCTGCAAAAGGTGAAAAGCACCAAGAGGTTTTCAGATGGAAGTGAAG
GTGACAGTGTGCTGGCAGCCCTCACAGCCCTCGCTTGTCTCCTTGGCGCTCTGCTGGGCTCCCACTTTGGCA
GCACTTGAGGAGCCCTTCAACCCGCGCTGCACTGTAGGAGCCCTTTCTGGGCTGGCCAAGGCCGGAGCCAGCT
CCCTCAGCTTGGGGGAGGTGCGGAGGAGAGGGGCGGGCAGGAACCGGGGCTGCGCGCAGCGCTTGGGGGCCAG
AGTGAGTTCGGGTGGGCGTGGGCTCGCGGGGGCCCCACTCAGAGCAGTGGCCGGCCCCAGGAGTGAGGGCT
TAGCACCTGGGCCAGCAGTGTGTGCTCGATTTCTCGCTGGGCCTTAGCTGCCTCCCCGCGGGGAGGGCTCGG
GACCTGCAGCCCTCCATGCTGACCTCCCCCACCCTCGTGGGCTCCTGTGCGGCGGAGCTCCCCAAGGAG
CGCCGCCCCCTGCTCCACAGCGCCAGTCCCATCGACCACCAAGGGCTGAGGAGTGGGGTGCACAGCGCGGGA
CTGGCAGGCAGTCCACCTGCTGCCCCAGTGTGGATCCACTGGGTGAAGCCAGCTGGGCTCCTGAGTCTGGTGG
GGACTTGGAGAACCTTTATGTCTAGCTAAGGGATTGTAATACACCGATGGGCACTCTGTATCTAGCTCAAGGTT
TGTAACACACCAATCAGCACCTGTGTCTAGCTCAGTGTGTGTAATGCACCAATCCACACTCTGTATCTGGCT
ACTCTGGTGGGACTTGAGAACCTTTGTGTCCACACTCTGTATCTAGCTAATCTAGTGGGGATGTGGGAACT
TTGTGTCTAGCTCAGGGATCGTAAACGCACCAATCAGCACCTGTCAAAACAGACCACTTGAATCTGTAAAT
GGACCAATCAGCAGGATGTGGGTGGGGCGAGACAAGAGAATAAAAGCAGGCTGCCTGAGCCAGCAGTGACAACCC
CCCTCGGGTCCCTCCACGCGCTGGAAGCTTTGTTCTTTCGCTCTTTGCAATAAATCTTGTACTGCCAAAA

542/550

FIGURE 542

MERAVRVESGVLVGVVCLLLACPATATGPEVAQPEVDTTLGRVRGRQVGVKGTDRLVNVFLGI
PFAQPPLGPD RFSAPHPAQPWEGVRDASTAPPMCLQDVESMNSSRFVLNGKQQIFSVSEDCLV
LNVYSPAIEVPAGSGRPVMVWVHGGALITGAATSYDGSALAAYGDVVVVTVQYRLGVLGFFSTG
DEHAPGNQGFLDVVAALRWVQENIAPFGGDLNCVTVFSGSAGGSIISGLVLSPVAAGLFHRAI
TQSGVITTPGIIDSHPWPLAQKIAN TLACSSSSPAEMVQCLQQKEGEELVLSKKLKNTIYPLT
VDGTVFPKSPKELLKEKPFHSVPFLMGVNNHEFSWLI PRGWGLLDTMEQMSREDMLAISTPVL
TSLDVPPPEMMPTVIDEYLG SNSDAQAKCQAFQEFMGDVFINVPTVSFSRYLRDSGSPVFFYEF
QHRPSSFAKIKPAWVKADHGAEGAFVFGGPFLMDESSRLAFPEATEEEKQLSLTMMAQWTHFA
RTGDPNSKALPPWPQFNQAEQYLEINPVPRAGQKFREAWMQFWSETLPSKIQQWHQKQKNRKA
QEDL

Important features:**Signal peptide:**

amino acids 1-27

Transmembrane domain:

amino acids 226-245

N-glycosylation site.

amino acids 105-109

N-myristoylation sites.

amino acids 10-16, 49-55, 62-68, 86-92, 150-156, 155-161,
162-168, 217-223, 227-233, 228-234, 232-238, 262-268, 357-363,
461-467

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 12-23

Carboxylesterases type-B serine active site.

amino acids 216-232

543/550

FIGURE 543

TGTCGCTGGCCCTCGCCATGACAGACCCCGGAGCGTCCCCTCCCCGCGCCCTCCTGCTTCTGCTGCTGCTA
CTGGGGGGCGCCACGGCCTCTTTCCTGAGGAGCCGCGCGCTTAGCGTGCGCCCCAGGGACTACCTGAACCAC
TATCCCGTGTTTGTGGGCAGCGGGCCCGGACGCTGACCCCGCAGAAGGTGCTGACGACCTCAACATCCAGCGA
GTCTGCGGGTCAACAGGACGCTGTTTATTGGGGACAGGGACAACCTCTACCGCGTAGAGCTGGAGCCCCCAGC
TCCACGGAGCTGCGGTACCAGAGGAAGCTGACCTGGAGATCTAACCCAGCGACATAAACGTGTGTGCGATGAAG
GGCAAACAGGAGGGCGAGTGTGAAACTTCGTAAAGGTGCTGCTCCTTCGGGACGAGTCCACGCTCTTTGTGTGC
GGTTCCAACGCTTCAACCCGGTGTGCGCCAACCTACAGCATAGACACCTGCAGCCCGTCGGAGACAACATCAGC
GGTATGGCCCGCTGCCGTACGACCCCAAGCAGCAGCAATGTTGCCCTCTTCTCTGACGGGATGCTCTTTCACAGCT
ACTGTTACCGACTTCTAGCCATTGATGCTGTCTATCTACCGCAGCCTCGGGGACAGGCCCACCTGCGCACCGTG
AAACATGACTCCAAGTGGTTCAAGAGCCTTACTTTGTCCATGCGGTGGAGTGGGGCAGCCATGTCTACTTCTTC
TTCCGGGAGATTGCGATGGAGTTAACTACCTGGAGAAGGTGGTGGTGTCCCGCGTGGCCGAGTGTGCAAGAAC
GACGTGGGAGGCTCCCCCGCGTGTGGAGAAGCAGTGGACGTCCTTCTGAAGGCGCGGCTCAACTGCTCTGTA
CCCGGAGACTCCCATTCTACTTCAACGTGCTGACGGCTGTACAGGGCGTGGTGAGCCTCGGGGGCCGGCCCGTG
GTCTGGCCGTTTTTTCCACGCCAGCAACAGCATCCCTGGCTCGGCTGTCTGCGCCTTTGACCTGACACAGGTG
GCAGCTGTGTTTGAAGGCCGCTTCGAGAGCAGAAGTCCCCCGAGTCCATCTGGACGCCGCTGCCGGAGGATCAG
GTGCCTCGACCCCGCCCGGGTGTGCGCAGCCCCCGGGATGCAGTACAATGCCTCCAGCGCCTTGCCGGATGAC
ATCCTCAACTTTGTCAAGACCCACCTCTGATGGACGAGGCGGTGCCCTCGCTGGGCCATGCGCCCTGGATCCTG
CGGACCTGATGAGGCACAGCTGACTCGAGTGGCTGTGGACGTGGGAGCCGCGCCCTGGGGCAACAGACCGTT
GTCTTCTGGGTCTGAGGCGGGGACGGTCTCAAGTCTCTCGTCCGGCCCAATGCCAGCACCTCAGGGACGTCT
GGGCTCAGTGTCTTCTGAGGAGTTTGTAGACCTACCGCCGAGCAGGTGTGGACGGCCCGCGGTGGCGAGACA
GGGACGGCTGCTGAGCTTGGAGCTGGACGCAGCTTCGGGGGGCCTGCTGGCTGCCTTCCCCCGCTGCGTGGTC
CGAGTGCCTGTGGCTCGCTGCCAGCAGTACTCGGGGTGTATGAAGAACTGTATCGGCAGTCAGGACCCCTACTGC
GGGTGGGCCCCCGACGGTCTTCCCTGACATCTTCTCAGCCCGGGCACCAGAGCCGCTTTGAGCAGGACGTGTCCGG
GCCAGCACCTCAGGCTTAGGGGACTGCACAGGACTCCTGCGGGCCAGCCTCTCCGAGGACCGCGCGGGGCTGGTG
TCGGTGAACCTGCTGGTAACGTGCTCGGTGGCGGCCCTTCGTGGTGGGAGCCGTGGTGTCCGGCTTACGCGTGGGC
TGGTTCTGGGCTCCGTGAGCGCGGGGAGCTGGCCCGGCGCAAGGACAAGGAGGCCATCCTGGCGCACGGGGCG
GGCAGGCGGTGCTGAGCGCTCAGCCGCTGGCGAGCGCAGGGTCCCGGGGGCGGGGCGGAGGCGGT
GGCGGTGGCGCGGGGTTCCCGGAGGCCCTGCTGGCGCCCTGATGCAGAACGGCTGGGCCAAGGCCACGCTG
CTGCAGGGCGGGCCCCACGACCTGGACTCGGGGTGCTGCCACGCCGAGCAGACGCCGCTGCCGAGAAGCGC
CTGCCCACCTCCGACCCGACCCCCACGCCCTGGGCCCCCGCGCCTGGGACCACGGCCACCCCTGCTCCCGCC
TCCGCTTCTATCTCCTCCTGCTGCTGGCGCCCGCCGGGCCCCCGAGCAGCCCCCGCGCTGGGGAGCCGACC
CCGACGGCCGCTCTATGCTGCCCGCCCGCGCGCCTCCACGGCGACTTCCGCTCACCCCCACGCCAGC
CCGACCGCCGCGGGTGGTGTCCGCGCCACGGGCCCTTGGACCCAGCCTCAGCCGCGATGGCCTCCCGCGG
CCCTGGAGCCCGCCCCGACGGGCAGCCTGAGGAGGCCACTGGGCCCCACGCCCTCCGGCCGCCACCTGCGC
CGCACCCACAGTTCAACAGCGCGGAGGCCCGGCTGGGGACCGCCACCGCGGTGCCACGCCCGCGCGGGCACA
GACTTGGCCACCTCCTCCCCATAGGGGGGGCGGACAGGACTGCGCCCCCGTGCCCTAGGCGGGGGCCCCCG
ATGCTTGGCAGTGCCAGCCACGGGAACAGGAGCGAGAGCGGTGCCAGAACGCCGGGGCCCGGGCAACTCCG
AGTGGGTGCTCAAGTCCCCCGCGACCCACCCGCGAGTGGGGGGCCCCCTCCGCCACAAGGAAGCACACCAG
CTCGCCCTCCCCCTACCCGGGCGCGAGGACGCTGAGACGTTTGGGGGTGGGTGGGCGGAGGACTTTGCTATG
GATTTGAGTTGACCTTATGCGGTAGGTTTTGGTTTTTTTTTGCAGTTTTGGTTTTCTTTGCGTTTTCTAAC
AATTGCACAACCTCGTTCTCGGGGTGGCGGCAGGAGGGGAGGCTTGGACGCCGCTGGGGAATGGGGGGCCACAG
CTGCAGACCTAAGCCCTCCCCACCCCTGGAAAGGTCCCTCCCCAACCCAGGCCCTGGCGTGTGTGGGTGTGCG
TGCGTGTGCGTCCGTTGTTGTCGTAAGGGGCCGGGAGGTGGGCGTGTGTGTGCGTGCCAGCGAAGGCTGCTG
TGGGCGTGTGTGTAAGTGGGCCACGCTGCAGGCTGTGTGTCCACGAGCGACGATCGTGGTGGCCCCAGCGGCC
TGGGCGTGTGGCTGAGCCGACGCTGGGGCTTCCAGAAGGCCCGGGGTCTCCGAGGTGCCGGTTAGGAGTTTGAAC
CCCCCCCCTCTGCAGAGGGAAGCGGGGACAATGCCGGGGTTTTCAGGCAGGAGACACGAGGAGGGCTGCCCGGA
AGTCACATCGGCAGCAGCTGTCTAAAGGGCTTGGGGGCGGGGCGGGAAGGTGGGTGGGGCCCCCTGTGTA
ATACGGCCCCAGGGTGGTGAAGAGTCCATGCCACCCGTCCTTGTGACCTCCCCCTATGACCTCCAGCTGA
CCATGCATGCCACGTGGCTGGCTGGGTCTCTGCCCTCTTTGGAGTTTGCCTCCCCAGCCCCCTCCCCATCAAT
AAAACCTCTGTTTACAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

544/550

FIGURE 544

MQTPRASPPRPALLLLLLLLGGAGHLFPEEPPPLSVAPRDYLNHYVPVFGSGPGRLTPAEGAD
DLNIQVRVLRVNRTLFIGDRDNLRYVELEPPTSTELRYQRKLTWRSNPSDINVCRMKGKQEGEC
RNFVKVLLLRDESTLFVCGSNAFNPVCANYSIDTLQPVGDNISGMARCPYDPKHANVALFSDG
MLFTATVTDFLAIDAVIYRSLGDRPTLRTVKHDSKWFEKPYFVHAVEWGSHVYFFFREIAMEF
NYLEKVVVSRVARVCKNDVGGSPRVLEKQWTSFLKARLNCSVPGDSHFYFNVLQAVTGVVSLG
GRPVVLAVFSTPSNSIPGSAVCAFDLTQVAAVFEGRFREQKSPESIWTPVPEDQVPRPRPGCC
AAPGMQYNASSALPDDILNFVKTHPLMDEAVPSLGHPWILRTLMRHQLTRVAVDVGAGPWGN
QTVVFLGSEAGTVLKFLVRPNASTSGTSGLSVFLEEFETYRPDRCGRPGGETGQRLLSLELD
AASGGLLAAPRCVVRVPVARCQQYSGCMKNCIGSQDPYCGWAPDGSCIFLSPGTRAAFEQDV
SGASTSGLGDCTGLLRASLSEDRAGLVSVNLLVTSSVAAFVVGAVVSGFSVGWVGLRERREL
ARRKDKEAILAHGAGEAVLSVSRLEGERRAQGPGGGGGGGGAGVPPPEALLAPLMQNGWAKAT
LLQGGPHDLDSGLLPTPEQTPLPQKRLPTPHPHALGPRAWDHGHPLLPASASSLLLLLAPA
RAPEQPPAPGEPTPDGRLYAARPGRASHGDFLTPHASPDRRRVVSAPTGPLDPASAADGLPR
PWSPPPTGSLRRPLGPHAPPAATLRRTHTFNSGEARPGDRHRGCHARPGTDLAHLPLYGGADR
TAPPVP

Important features:**Signal peptide:**

amino acids 1-25

Transmembrane domains:

amino acids 318-339, 598-617

N-glycosylation sites.amino acids 74-78, 155-159, 167-171, 291-295, 386-390, 441-445,
462-466**Glycosaminoglycan attachment sites.**

amino acids 51-55, 573-577

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation sites.amino acids 21-27, 50-56, 189-195, 333-339, 382-388, 448-454,
490-496, 491-497, 508-514, 509-515, 531-537, 558-564, 569-575,
574-580, 580-586, 610-616, 643-649, 663-669, 666-672, 667-673,
668-674, 669-675, 670-676, 868-874, 879-885

545/550

FIGURE 545

GATGGCGCAGCCACAGCTTCTGTGAGATTGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGGG
GACCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTAT**ATG**CGTCAATTCCCCAAA
ACAAGTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTT
CCAGGCCTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACG
ACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAG
GCTTGGCAGTTTTTCTTACTCCTGTGGTCTCCAGATTTCAGGCCTAAGATGAAAGCCTCTAGT
CTTGCCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAG
ACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCT
GAGATACGGGGCAGTGTGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACT
GAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTC
TATCTGGACAGGGTATTTAAAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGC
AGCCTCGCCAATTCCTTTCTTACCATCAAGAAGGACCTCCGGCTCTCTCATGCCCACATGACA
TGCCATTGTGGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTG
GAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGGAAGTAGACATTCTTCTGCAATGGATGGAG
GAGACAGAA**TAG**GAGGAAAGTGATGCTGCTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCT
TCAATACCTGCAGAGGAGGCATGACCCCAAACCACCATCTCTTTACTGTACTAGTCTTGTGCT
GGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCATGATTGTCTTTATGCATCCCC
AATCTTAATTGAGACCATACTTGTATAAGATTTTTGTAATATCTTTCTGCTATTGGATATATT
TATTAGTTAATATATTTATTTATTTTTTGTCTATTTAATGTATTTATTTTTTTACTTGGACATG
AACTTTAAAAAAATTACAGATTATATTTATAACCTGACTAGAGCAGGTGATGTATTTTTAT
ACAGTAAAAAATAAACCTTGTAATTCTAGAAGAGTGGCTAGGGGGGTATTCAATTTGTAT
TCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGATATTTGAAATTGAACCAATGAC
TACTTAGGATGGGTGTGGAATAAGTTTTGATGTGGAATTGCACATCTACCTTACAATTACTG
ACCATCCCCAGTAGACTCCCCAGTCCCATAAATTGTGTATCTTCCAGCCAGGAATCCTACACGG
CCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATACCAAAAAAAAAAAAAAAAAAAAA

546/550

FIGURE 546

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPE
IFSSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQE
IRNGFSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHY
TLRKISSLANSFLLTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDI
LLQWMEETE

Important features:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

547/550

FIGURE 547

AGCAACTCAAGTTCATCATTGTCCTGAGAGAGAGGAGCAGCGCGGTTCTCGGCCGGGACAGCA
GAACGCCAGGGGACCCTCACCTGGGCGCGCCGGGGCACGGGCTTTGATTGTCCTGGGGTCGCG
GAGACCCGCGCGCCTGCCCTGCACGCCGGGCGGCAACCTTTGCAGTCGCGTTGGCTGCTGCGA
TCGGCCGGCGGGTCCCTGCCGAAGGCTCGGCTGCTTCTGTCCACCTCTTACACTTCTTCATTT
ATCGGTGGATCATTTTCGAGAGTCCGTCTTGTAAATGTTTGGCACTTTGCTACTTTATTGCTTC
TTTCTGGCGACAGTTCCAGCACTCGCCGAGACCGGCGGAGAAAGGCAGCTGAGCCCGGAGAAG
AGCGAAATATGGGGACCCGGGCTAAAAGCAGACGTCGTCCTTCCCGCCCGCTATTTCTATATT
CAGGCAGTGGATACATCAGGGAATAAATTCACATCTTCTCCAGGCGAAAAGGTCTTCCAGGTG
AAAGTCTCAGCACCAGAGGAGCAATTCAGTAGAGTTGGAGTCCAGGTTTTAGACCGAAAAGAT
GGGTCCTTCATAGTAAGATACAGAATGTATGCAAGCTACAAAAATCTGAAGGTGGAATTAAA
TTCCAAGGGCAACATGTGGCCAAATCCCATATATTTTAAAAGGGCCGGTTTACCATGAGAAC
TGTGACTGTCCTCTGCAAGATAGTGCAGCCTGGCTACGGGAGATGAACTGCCCTGAAACCATT
GCTCAGATTTCAGAGAGATCTGGCACATTTCCCTGCTGTGGATCCAGAAAAGATTGCAGTAGAA
ATCCCAAAAAGATTTGGACAGAGGCAGAGCCTATGTCACTACACCTTAAAGGATAACAAGGTT
TATATCAAGACTCATGGTGAACATGTAGGTTTTAGAAATTTTCATGGATGCCATACTACTTTCT
TTGACTAGAAAAGGTGAAGATGCCAGATGTGGAGCTCTTTGTTAATTTGGGAGACTGGCCTTTG
GAAAAAAGAAATCCAATTCAAACATCCATCCGATCTTTTCTGGTGTGGCTCCACAGATTCC
AAGGATATCGTGATGCCTACGTACGATTTGACTGATTCTGTTCTGGAACCATGGGCCGGGTA
AGTCTGGATATGATGTCCGTGCAAGCTAACACGGGTCCTCCCTGGGAAAGCAAAAATTCCTACT
GCCGTCTGGAGAGGGCGAGACAGCCGCAAAGAGAGACTCGAGCTGGTTAACTCAGTAGAAAA
CACCCAGAACTCATAGACGCTGCTTTCACCAACTTTTTCTTCTTTAAACACGATGAAAACCTG
TATGGTCCCATTTGTGAAACATATTTCAATTTTTTGATTTCTTCAAGCATAAGTATCAAATAAAT
ATCGATGGCACTGTAGCAGCTTATCGCCTGCCATATTTGCTAGTTGGTGACAGTGTTGTGCTG
AAGCAGGATTCCATCTACTATGAACATTTTTTACAATGAGCTGCAGCCCTGGAAACACTACATT
CCAGTTAAGAGCAACCTGAGCGATCTGCTAGAAAACTTAAATGGGCGAAAGATCACGATGAA
GAGGCCAAAAGATAGCAAAAGCAGGACAAGAATTTGCAAGAAATAATCTCATGGGCGATGAC
ATATTCTGTTATTATTTCAAACCTTTCCAGGAATATGCCAATTTACAAGTGAGTGAGCCCCAA
ATCCGAGAGGGCATGAAAAGGGTAGAACCACAGACTGAGGACGACCTCTTCCCTTGTAATTGC
CATAGGAAAAAGACCAAAGATGAACTCTGATATGCAAAATAACTTCTATTAGAATAATGGTGC
TCTGAAGACTCTTCTTAACTAAAAGAAGAATTTTTTTAAGTATTAATTCCATGGACAATATA
AAATCTGTGTGATTGTTTGCAGTATGAAGACACATTTCTACTTATGCAGTATTCTCATGACTG
TACTTTAAAGTACATTTTTAGAAATTTTATAATAAAACCACCTTTATTTTAAAGGAAAAAA

548/550

FIGURE 548

MFGTLLLYCFFLATVPALAETGGERQLSPEKSEIWGPGLKADVLPARYFYIQAVDTSGNKFT
SSPGEKVFQVKVSAPEEQFTRVGVQVLDRKDGSFIVRYRMYASYKNLKVEIKFQGQHVAKSPY
ILKGPVYHENCDCPLQDSAAWLREMNCPETIAQIQRDLAHFPAVDPEKIAVEIPKRFGQRQSL
CHYTLKDNKVYIKTHGEHVGFRI FMDAILLSLTRKVKMPDVELFVN LGDWPLEKKKSNSNIHP
IFSWCGSTDSKDIVMPTYDLTDSVLETMGRVSLDMSVQANTGPPWESKNSTAVWRGRDSRKE
RLELVKLSRKHP ELIDAAFTNFFFFKH DENLYGPIVKHISFFDFFKH KYQINIDGTVAAYRLP
YLLVGDSVVLKQDSIYYEHFYNELQPWKHYIPVKS NLSDLLEK LKWAKDHDEEAKKIAKAGQE
FARNNLMGDDIFCYF KLFQ EYANLQVSEPQIREGMKRVEPQTEDDLFPCTCHRKKTKDEL

Important features:**Signal peptide:**

amino acids 1-17

N-glycosylation sites.

amino acids 302-306, 414-418

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 243-247, 495-499

Tyrosine kinase phosphorylation site.

amino acids 341-348

N-myristoylation sites.

amino acids 59-65, 118-124, 184-190, 258-264, 370-376, 439-445

Endoplasmic reticulum targeting sequence.

amino acids 499-504

549/550

FIGURE 549

GGGTGATTGAACTAAACCTTCGCCGCACCGAGTTTGCAGTACGGCCGTCACCCGCACCGCTGC
CTGCTTGCGGTTGGAGAAATCAAGGCCCTACCGGGCCTCCGTAGTCACCTCTCTATAGTGGGC
GTGGCCGAGGCCGGGGTGACCCTGCCGGAGCCTCCGCTGCCAGCGAC**ATG**TTCAAGGTAATTC
AGAGGTCCGTGGGGCCAGCCAGCCTGAGCTTGCTCACCTTCAAAGTCTATGCAGCACCAAAAA
AGGACTCACCTCCCAAAAATTCGTTGAAGGTTGATGAGCTTTCACCTACTCAGTTCCTGAGG
GTCAATCGAAGTATGTGGAGGAGGCAAGGAGCCAGCTTGAAGAAAGCATCTCACAGCTCCGAC
ACTATTGCGAGCCATACACAACCTGGTGTGTCAGGAAACGTACTCCCAAACTAAGCCCAAGATGC
AAAGTTTGGTTCAATGGGGGTTAGACAGCTATGACTATCTCCAAAATGCACCTCCTGGATTTT
TTCCGAGACTTGGTGTATTGGTTTTGCTGGCCTTATTGGACTCCTTTTGGCTAGAGGTTCAA
AAATAAAGAAGCTAGTGTATCCGCCTGGTTTCATGGGATTAGCTGCCTCCCTCTATTATCCAC
ACAAGCCATCGTGTGTTGCCAGGTCAGTGGGGAGAGATTATATGACTGGGGTTTACGAGGAT
ATATAGTCATAGAAGATTTGTGGAAGGAGAACTTTCAAAAGCCAGGAAATGTGAAGAATTCAC
CTGGAATAAG**TAG**AAAACTCCATGCTCTGCCATCTTAATCAGTTATAGGTAAACATTGGAAA
CTCCATAGAATAAATCAGTATTTCTACAGAAAAATGGCATAGAAGTCAGTATTGAATGTATTA
AATTGGCTTTCTTCTTCAGGAAAACTAGACCAGACCTCTGTTATCTTCTGTGAAATCATCCT
ACAAGCAAATAACCTGGAATCCCTTCACCTAGAGATAATGTACAAGCCTTAGAACTCCTCAT
TCTCATGTTGCTATTTATGTACCTAATTAAACCCAAGTTTAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAA

550/550

FIGURE 550

MFKVIQRSVGPASLSLLTFKVYAAPKKDSPPKNSVKVDELSTLYSVPEGQSKYVEEARSQLEES
ISQLRHYCEPYTTWCQETYSQTKPKMQSLVQWGLDSYDYLQNAPPGFFPRLGVIGFAGLIGLL
LARGSKIKKLVYPPGFMGLAASLYYPQQAIVFAQVSGERLYDWGLRGYIVIEDLWKENFQKPG
NVKNSPGTK

Important features:

Signal peptide:

Amino acids 1-23

Transmembrane domain:

Amino acids 111-130

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 26-30

Tyrosine kinase phosphorylation site:

Amino acids 36-44

N-myristoylation sites:

Amino acids 124-130;144-150;189-195



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

